Fuzzy Sets and Systems 542 (2026) 109958

Contents lists available at ScienceDirect

Fuzzy Sets and Systems

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

An arctangent entropy decision tree based on explicit hesitation and its application in thrombus classification

Xiaozeng Xu [a,b], Yanni Guo [a], Xing Pu [a], Weihua Xu [c,*]

[a] College of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing, 400067, China
[b] Chongqing Key Laboratory of Statistical Intelligent Computing and Monitoring, Chongqing, 400067, China
[c] College of Artificial Intelligence, Southwest University, Chongqing, 400715, China

ARTICLE INFO

Keywords:
Intuitionistic fuzzy decision tree
Arctangent intuitionistic fuzzy entropy
Fuzzy clustering
Classification prediction
Classification of thrombus

ABSTRACT

The accurate classification of thrombus is crucial for clinical diagnosis and treatment. However, due to its complex pathogenesis, the high dimensionality of relevant features and inherent uncertainties, constructing robust classification models presents numerous difficulties. To address the challenge of handling fuzzy and uncertain information in such complex data, this study introduces a novel explicit arctangent entropy that incorporates hesitation, where the entropy value increases monotonically with the level of hesitation. It is theoretically established that the proposed entropy satisfies the axiomatic requirements of intuitionistic fuzzy entropy. Based on this, an intuitionistic fuzzy decision tree with an explicit hesitation model was further constructed. Empirical analysis across multiple public datasets demonstrates that the proposed model outperforms other decision tree models combining fuzzy entropies in terms of Accuracy, F1-Score, and AUC. It also demonstrates strong competitiveness in comprehensive comparisons with seven mainstream classifiers such as C4.5, RF, and XGBoost. Finally, the model was applied to the clinical thrombus classification task. On clinical medical data, the model achieved outstanding Accuracy (0.9032), validating its effectiveness and practical value in classifying complex medical data.

1. Introduction

Thrombus [1,2] as a common cardiovascular disease requires early recognition and precise diagnosis. However, clinical thrombus classification tasks are constrained by complex correlations among high-dimensional features, the inherent ambiguity of medical indicators, and inconsistent data quality, resulting in limited discriminative performance of traditional classification models in practical applications. Therefore, developing classification models capable of effectively handling such uncertainties has become a research direction of significant clinical importance.

Abbreviations: AIFE, Arctangent Intuitionistic Fuzzy Entropy with Explicit Hesitation; AIFE-IIFDT, Intuitionistic Fuzzy Decision Tree with Explicit Hesitation; C4.5, C4.5 Decision Tree; CART, Classification and Regression Tree; COPRAS, Complex Proportional Assessment; E_{AIFE}, Proposed a novel intuitionistic fuzzy entropy; E_{JQ}, Jiang et al.'s Gaussian Curvature-based Entropy; E_{KS}, Kashyap et al.'s Secant-based Entropy; E_{LS}, Lan's Composite Cosine-Exponential-Logarithmic Entropy; E_{LY}, Li's Intuitionistic Fuzzy Entropy; E_{XL}, Xu and Li's Arctangent Intuitionistic Fuzzy Entropy; FCM, Fuzzy C-Means Clustering; GBDT, Gradient Boosting Decision Tree; GCF, Gaussian Curvature Filter; ID3, Iterative Dichotomiser 3; IFDT, Ren et al.'s Intuitionistic Fuzzy Decision Tree; IFS, Intuitionistic Fuzzy Sets; IIFDT, Li's Improved Intuitionistic Fuzzy Decision Tree; RBFNN, Radial Basis Function Neural Network; RF, Random Forest; SLIQ, Supervised Learning in Quest; XGBoost, Extreme Gradient Boosting.

* Corresponding author.
E-mail addresses: xiaozeng.xu@ctbu.edu.cn (X. Xu), 940316038@qq.com (Y. Guo), pxpx1f@163.com (X. Pu), chxuwh@swu.edu.cn (W. Xu).

https://doi.org/10.1016/j.fss.2026.109958
Received 19 November 2025; Received in revised form 1 March 2026; Accepted 15 May 2026
Available online 23 May 2026

The challenges faced in the aforementioned classification of thrombus fundamentally stem from the core issue of handling ambiguous and uncertain information in data classification. Real-world datasets often contain noise [3], incompleteness, and inherent ambiguity, making the accurate handling of uncertain information a central challenge in improving classification performance. In recent years, intuitionistic fuzzy decision trees have demonstrated theoretical advantages in expressing complex uncertainties through the integration of intuitionistic fuzzy sets, while maintaining the exceptional interpretability inherent to decision tree models. This approach has significant theoretical research value and broad applications prospects for handling complex fuzzy classification problems.

The key contributions of this research are as follows.

This study proposes an AIFE entropy and its corresponding AIFE-IIFDT model to address fuzzy and uncertain classification problems.

(1) Propose a new axiomatic and discriminative fuzzy entropy. This entropy incorporates hesitation through a mathematical model, causing its value to increase monotonically with uncertainty. This enables precise differentiation of uncertainty among samples with identical membership differences but varying degrees of uncertainty.
(2) Construct a comprehensive classification framework centered on AIFE entropy as the core criterion for fragmentation. Integrating the proposed entropy into an intuitionistic fuzzy decision tree significantly enhances the model's ability to process complex fuzzy information.
(3) The effectiveness and practical value of the proposed model have been validated through extensive experimentation on multiple public datasets.

The structure of this work is as follows: Section 2 outlines the relevant theoretical background. Section 3 introduces the fundamental theories of intuitionistic fuzzy sets and intuitionistic fuzzy entropy, and details the complete construction process of the proposed AIFE entropy and AIFE-IIFDT models. Section 4 describes the experimental setup and reports experimental results and analyzes from multiple dimensions. Section 5 demonstrates the practical application of AIFE-IIFDT model in clinical thrombus classification tasks. Finally, Section 6 summarizes the entire work and outlines future research directions.

2. Literature survey

To address complex and uncertain problems, this section presents numerous studies from various perspectives to achieve the proposed goals. Section 2.1 reviews traditional decision trees. Section 2.2 focuses on the development of fuzzy decision trees. Section 2.3 discusses research on intuitionistic fuzzy entropies. Section 2.4 reviews intuitionistic fuzzy decision trees. Section 2.5 clarifies the motivation for this study.

2.1. Traditional decision trees

Decision trees have become one of the foundational models in classification tasks because of their clear structure and high interpretability. Quinlan [4] pioneered the ID3 algorithm for decision tree induction learning. Subsequently, Hssina et al. [5] conducted a systematic comparative study of the ID3 and C4.5 algorithms, advancing the refinement and development of the decision tree theory. The effectiveness of this approach has been validated in numerous domains. For example, Jiang et al. [6] applied it to dynamic trust evaluation in underwater wireless sensor networks. However, traditional decision trees rely on a hard-splitting mechanism that requires discretizing continuous attributes, leading to inevitable information loss. More fundamentally, this approach struggles to effectively capture data ambiguity and uncertainty, particularly when dealing with complex decision boundaries.

2.2. Fuzzy decision trees

Fuzzy decision trees address the limitations of traditional decision trees in processing uncertain information by incorporating fuzzy set theory, thereby achieving a soft partition of classification boundaries. Established by Zadeh [7], fuzzy set theory employs membership functions to describe partial membership relationships between elements and sets, providing a mathematical foundation for handling uncertain data. The theory and methods of fuzzy set entropy have also been developed accordingly [8]. For example, Smith et al. [9] proposed a conditional hybrid entropy, which demonstrated excellent performance in a case study on prostate cancer diagnosis, significantly outperforming traditional Shannon entropy and Zadeh's entropy. Janikow [10] provided the key theoretical foundations for the development of fuzzy decision trees. Umanol et al. [11] pioneered the integration of this concept into decision trees with the fuzzy ID3 algorithm. Subsequently, Chandra and Varghese [12,13] proposed fuzzy SLIQ decision trees and fuzzy Gini-index-based decision trees, significantly enhancing the models' robustness to noise and boundary samples. The practical value of these methods has been demonstrated through successful applications across various fields, including industrial process monitoring [14,15], medical diagnosis [16], anomaly detection [17,18], and signal processing [19,20]. Recent theoretical advances in this field are reflected in contributions such as the fuzzy belief decision tree proposed by Jiao et al. [21] and the decision tree based on a fuzzy rule regression model developed by Zhu et al. [22], which continue to drive theoretical innovation and expand practical applications.

2.3. Intuitionistic fuzzy entropy

Since Atanassov [23] proposed the theory of IFS, researchers have been dedicated to developing entropy functions that can effectively measure intuitionistic fuzzy uncertainty. As a core metric for measuring uncertainty, research on its measurement methods has

made a series of advances. Burillo and Bustince [24] provided an axiomatic definition of intuitionistic fuzzy entropy, while Szmidt and Kacprzyk [25] proposed a distance based on an entropy measurement method from a geometric perspective. Subsequently, various types of intuitionistic fuzzy entropy functions have been successively proposed and continuously refined, including entropy based on Gaussian curvature [26], entropy applicable to multi-criteria decision-making [27,28], cosine-based and cosine-exponential logarithmic intuitionistic fuzzy entropy [29], arctangent entropy [30]. In recent years, this field has shown a trend toward multidirectional convergence and integration. For instance, Aggarwal [31] has dedicated efforts to bridge the theoretical gap between probabilistic entropy and fuzzy entropy. Liu et al. [32] built an explainable classifier based on axiomatic fuzzy sets and semantic entropy, while Zhang et al. [33] explored an online streaming feature selection algorithm based on fuzzy Gini entropy.

2.4. Intuitionistic fuzzy decision trees

The theoretical development of intuitionistic fuzzy entropy has laid the foundation for its application in intuitionistic fuzzy decision trees. To overcome the limitation of fuzzy decision trees that depend solely on membership, the researchers introduced nonmembership and hesitation, thereby advancing the development of intuitionistic fuzzy decision trees. Intuitionistic fuzzy decision trees offer a more powerful solution for classifying complex and uncertain data by integrating the comprehensive expressive power of IFS with the interpretability advantages of decision trees. Research on intuitionistic fuzzy decision trees has demonstrated a comprehensive development from theoretical frameworks to practical applications. (a) In terms of model construction, Bujnowski et al. [34] pioneered the development of a framework for the construction of intuitionistic fuzzy decision trees. Subsequently, Li [35] further proposed the IIFDT, establishing a complete workflow from data handling to rule generation. Xian et al. [36] introduced a decision tree structure based on an intuitionistic fuzzy twin vector machine. (b) At the application level, such models have shown remarkable effectiveness in actual scenarios such as e-commerce fraud detection [37] and stroke diagnosis [38], demonstrating their practical value in handling complex uncertainty problems.

2.5. Motivation

Based on the literature review, complex and ambiguous classification tasks face numerous challenges.

(1) High dimensionality coexists with noise. clinical data often contain substantial redundant features and measurement noise, making it difficult for traditional models to effectively extract key distinguishing information.
(2) Three-dimensional uncertainty structures. Traditional models focus solely on membership, making it hard to fully characterize nonmembership and hesitation, resulting in incomplete uncertainty representation.
(3) A dual balance of accuracy and confidence. Particularly in medical diagnostics, the pursuit of classification accuracy while neglecting classification confidence can lead to misdiagnosis risks. Models must achieve high precision and high reliability.

In response to the challenges, existing approaches still have problems and have failed to develop a systematic solution. To provide a more direct comparison of the key differences between existing approaches, Table 1 systematically summarizes the methodological comparisons, advantages and disadvantages of various studies, laying the groundwork to clarify the motivation behind this research.

(1) Traditional decision trees. Employing a hard-cutting mechanism leads to information loss when discretizing continuous attributes. Furthermore, it cannot handle fuzzy relationships.
(2) Fuzzy decision trees. Relying only on the single dimension of membership, they neglect nonmembership and hesitation, thus failing to fully represent three-dimensional uncertainty.
(3) Existing intuitionistic fuzzy entropies. Some entropies underutilize uncertainty measures, resulting in limited discriminative capability. Some do not incorporate uncertainty, and some others cannot distinguish samples with the same difference in membership and nonmembership, but different hesitation levels.
(4) Existing intuitionistic fuzzy decision trees. Limitations in entropy function selection. Most intuitionistic fuzzy entropies are primarily applied in multi-attribute decision scenarios and have not yet been systematically integrated with intuitionistic fuzzy decision trees.

Therefore, the core motivation of this study is to focus on the key breakthrough of explicit use of uncertainty. By proposing AIFE entropy and its AIFE-IIFDT model, it addresses the limitations of current methods.

(1) The theoretical advantages of AIFE entropy. By introducing uncertainty as an independent variable, entropy values increase monotonically with hesitation, enabling precise measurement of uncertainty and providing reliable splitting criteria for decision trees.
(2) Construction of the AIFE-IIFDT model. Employing AIFE entropy as a splitting criterion, the decision tree actively balances membership variance and uncertainty during growth. This approach improves classification accuracy and robustness while preserving the interpretability of the decision tree.

Table 1
Tabular summary of existing classification methods and intuitionistic fuzzy entropy measures according to literature review.

Work	Methodology proposed	Advantages	Limitations
Decision Tree [4,5,39,40]	Hard splitting algorithms such as ID3, C4.5, and CART based on information gain or Gini index.	Clear structure, high interpretabil, and computational efficiency. C4.5 and CART support continuous attribute, with partial pruning introduced to mitigate overfitting.	Rigid partitioning mechanisms cannot handle fuzzy boundaries. They are completely incapable of expressing ambiguity and hesitation in data.
Fuzzy Decision Tree [11–13]	Implementing soft splitting through fuzzy sets, such as fuzzy ID3/SLIQ/Gini index decision tree.	More robust to noise and boundary samples, capable of handling numerical and symbolic features, with significantly reduced tree size.	Only membership degrees are considered, nonmembership degrees and uncertainty degrees are not introduced.
Ensemble Model [41–43]	RF, GBDT, XGBoost enhance generalization capabilities by integrating multiple decision trees.	Strong generalization capabilities, capable of evaluating feature importance, and achieving leading accuracy across classification tasks.	Model interpretability is reduced, as each tree remains a clear split, unable to explicitly model ambiguity.
Neural Network Classifier [44]	RBFNN utilize radial basis functions in hidden layers to achieve nonlinear mapping.	RBFNN possesses nonlinear mapping capabilities and exhibits high computational efficiency.	The model is a black box with poor interpretability and cannot handle hesitation in the data.
Intuitionistic Fuzzy Entropy Theory [25–27,29,30]	Measuring the uncertainty of fuzzy sets from geometric, Gaussian curvature, cotangent function, composite function, and arctangent perspectives.	It satisfies the axiomatic definition of entropy, quantifies fuzziness and uncertainty from multiple perspectives, and has demonstrated its validity in multi-attribute decision-making.	It remains largely confined to theoretical or decision-making levels, without integration into classifiers, and the role of hesitation in entropy functions is not explicitly defined.
Intuitionistic Fuzzy Decision Tree [35]	Fully implement IIFDT, including fuzzification, tree construction, pruning, and other operations.	Completely implement and validate the classification effectiveness.	The entropy function employed does not entirely utilize the uncertainty degree, incorporating it only in a linear manner.

3. Theoretical foundations and proposed methods

This chapter begins by systematically elaborating on the theoretical foundation of this study. Section 3.1 introduces the fundamental concepts of intuitionistic fuzzy entropy, while Section 3.2 reviews several major intuitionistic fuzzy entropies. Building upon this foundation, we propose a novel AIFE entropy. Section 3.3 provides the mathematical definition of this entropy and completes its axiomatic proof, establishing a rigorous theoretical foundation for node splitting. Subsequently, Section 3.4 outlines the complete data preprocessing workflow to provide the model with inputs that can be processed directly. Finally, Section 3.5 employs AIFE entropy as the splitting criterion to construct the complete AIFE-IIFDT model, with the overall workflow of this method shown in Fig. 1a. Additionally, Section 3.6 provides a detailed demonstration of the complete construction process of the proposed model through a sample calculation, while Section 3.7 conducts a theoretical analysis of the time complexity of the proposed method.

3.1. Intuitionistic fuzzy entropy

An intuitionistic fuzzy set extends traditional fuzzy set theory by incorporating both nonmembership and hesitation. Let x be a fixed domain. An intuitionistic fuzzy set A in x is defined as follows

$$A = \left\{ \langle x, \mu_A(x), \nu_A(x) \rangle \mid x \in X \right\} \tag{1}$$

where $\mu_A : X \to [0,1]$ and $\nu_A : X \to [0,1]$, such that $\forall x \in X$, satisfy

$$0 \leq \mu_A(x) + \nu_A(x) \leq 1 \tag{2}$$

The function $\mu_A(x)$ denotes the membership of element x belonging to set A, $\nu_A(x)$ represents the nonmembership of x belonging to A, and the hesitation is defined as

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x) \tag{3}$$

This value reflects the degree of uncertainty about whether x belongs to A, and satisfies $0 \leq \pi_A(x) \leq 1$. Intuitionistic fuzzy entropy serves as a crucial metric for quantifying the uncertainty in IFS. Its axiomatic definition was first introduced by Burillo and Bustince. Subsequent scholars have further refined this axiomatic framework from various perspectives. Continuously optimized the axiomatic definition of intuitionistic fuzzy entropy. Formally, a mapping $E : \mathrm{IFS}(X) \to [0,1]$ is called an intuitionistic fuzzy entropy if it satisfies the following conditions.

(1) $E(A) = 0$ if and only if A is a crisp set.
(2) $E(A) = 1$ if and only if $\mu_A(x_i) = \nu_A(x_i)$ holds for every $x_i \in X$.
(3) $E(A) = E(A^c)$.
(4) For every $x_i \in X$, $E(A) \geq E(B)$ if either of the following holds:
- when $\pi_A(x_i) = \pi_B(x_i)$ and $f_A(x_i) \leq f_B(x_i)$;
- or when $f_A(x_i) = f_B(x_i)$ and $\pi_A(x_i) \geq \pi_B(x_i)$.

3.2. Several types of intuitionistic fuzzy entropy

This section presents five representative types of intuitionistic fuzzy entropy(All serve as the baseline methods for the comparative experiments in this study), encompassing diverse mathematical structures such as trigonometric functions and composite functions to accommodate various application scenarios. The mathematical expressions for each entropy are provided below, along with brief descriptions of their core characteristics and typical application scenarios.

(1) Li [35] proposed the intuitionistic fuzzy entropy

$$E_{LY}(A) = \frac{1}{n}\sum_{i=1}^{n}\frac{1 - \left|\mu_A(x_i) - \nu_A(x_i)\right| + \pi_A(x_i)}{2} \tag{4}$$

- Core feature. Corrects the deficiency in Szmidt entropy that disregards hesitation, comprehensively considering both ambiguity and uncertainty.
- Application scenarios. Fuzzy decision tree attribute selection and node splitting, fuzzy classification, and prediction of structured data.

(2) $E_{JQ}(A)$ is an intuitionistic fuzzy entropy introduced by Jiang et al. [26], defined as

$$E_{JQ}(A) = 1 - [\frac{1}{3}|(3 - 2\pi_A(x_i))| \times |\mu_A(x_i) - \nu_A(x_i)|] \tag{5}$$

- Core feature. Integrates GCF with intuitionistic fuzzy entropy, combining entropy calculation with local structural information in images to effectively preserve critical details such as edges and textures.
- Application scenarios. Multimodal medical image fusion, blurred image processing, and feature extraction requiring preservation of local structural information.

(3) The form of intuitionistic fuzzy entropy $E_{KS}(A)$ presented by Kashyap et al. [27] is as follows

$$E_{KS}(A) = \frac{1}{n}\sum_{i=1}^{n}\frac{\sqrt{3}}{4-3\sqrt{3}}\left(\sec\frac{(1+\mu_A(x_i)-\nu_A(x_i))\pi}{6} + \sec\frac{(1+\nu_A(x_i)-\mu_A(x_i))\pi}{6} - 3\right) \tag{6}$$

- Core feature. Constructed around the cotangent trigonometric function, incorporates the nonlinear characteristics of trigonometric functions to measure the uncertainty difference between membership and nonmembership.
- Application scenarios. Multi-criteria decision-making, qualitative indicator ranking and fuzzy evaluation for low-dimensional, sparse data.

(4) Lan [29] developed an intuitionistic fuzzy entropy $E_{LS}(A)$, whose expression is

$$E_{LS}(A) = \frac{1}{2n}\sum_{i=1}^{n}\left[\cos\frac{|\mu_A(x_i) - \nu_A(x_i)|\, e^{-\pi_A(x_i)}\pi}{2} + \left(1 - |\mu_A(x_i) - \nu_A(x_i)|\right)\ln\left(\pi_A(x_i) + e - 1\right)\right] \tag{7}$$

- Core feature. Integrates characteristics of cosine, exponential, and logarithmic functions to nonlinearly capture multidimensional uncertainty. Exhibits strong discrimination capabilities with minimal parameters and intuitive interpretability.
- Application scenarios. Multi-criteria decision-making, fuzzy ranking and optimization of mixed qualitative and quantitative indicators.

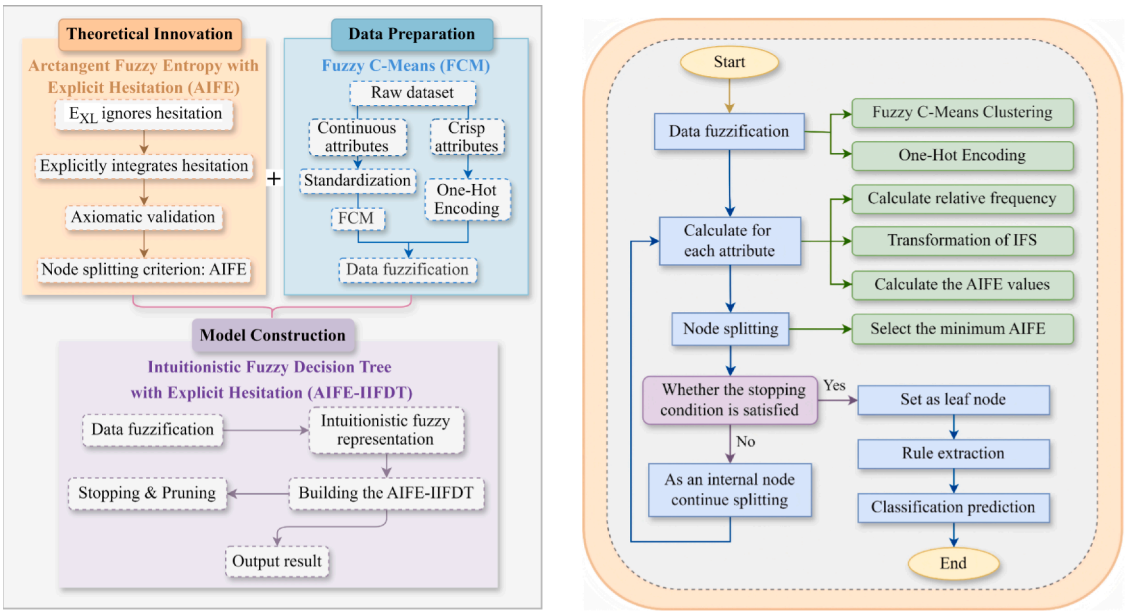
(5) Xu and Li [30] formulated an arctangent intuitionistic fuzzy entropy $E_{XL}(A)$, defined as follows

$$E_{XL}(A) = \frac{\frac{1}{n}\sum_{i=1}^{n}\left\{\frac{4}{\pi}\arctan\left(\frac{1+\mu_A(x_i)-\nu_A(x_i)}{2}\right) + \frac{4}{\pi}\arctan\left(\frac{1-\mu_A(x_i)+\nu_A(x_i)}{2}\right) - 1\right\}}{\frac{8}{\pi}\arctan\left(\frac{1}{2}\right) - 1} \tag{8}$$

- Core feature. Constructed based on the arctangent function, with entropy values exhibiting monotonic variation within the $[0, 1]$ interval. Then nonlinearly characterizes the deviation between membership and nonmembership, delivering high evaluative value distinctiveness.
- Application scenarios. Multi-attribute group decision-making, fuzzy evaluation and ranking, simple fuzzy classification of low-dimensional data.

3.3. Arctangent intuitionistic fuzzy entropy with explicit hesitation

Intuitionistic fuzzy entropy is the core for building intuitionistic fuzzy decision trees. This section proposes a new AIFE entropy, which is introduced in three distinct parts.



(a) Overall framework diagram (b) Flowchart of the AIFE-IIFDT

Fig. 1. Overall framework and workflow of the AIFE-IIFDT model. (a) Three core components of the AIFE-IIFDT model. (b) Detailed flowchart for constructing the AIFE-IIFDT model.

3.3.1. Design motivation

The arctangent intuitionistic fuzzy entropy introduced by Xu and Li [30] in 2023 exhibits desirable nonlinear properties and good intuitive consistency, making it effective in multi-attribute decision-making contexts. However, this entropy primarily measures the divergence between membership and nonmembership, without explicitly modeling the role of hesitation in uncertainty.

To illustrate this limitation concretely, consider $a_1, a_2 \in \mathrm{IFSs}(X)$, $a_1 = (\mu = 0.5, \nu = 0.3)$, $a_2 = (\mu = 0.6, \nu = 0.4)$, calculating intuitionistic fuzzy entropy. Both have the same $\mu - \nu = 0.2$. Intuitively, a_1 exhibits greater uncertainty due to its 20% degree of hesitation. However, when using E_{XL} (Eq. (8)), since its calculation relies solely on the combination of $(1 + \mu - \nu)/2$ and $(1 - \mu + \nu)/2$, this results in $E_{XL}(a_1) = E_{XL}(a_2)$. This indicates that E_{XL} cannot distinguish between uncertainties arising from the same difference in membership and nonmembership but differing hesitation, rendering its uncertainty measurement incomplete.

To solve this question, this study introduces a correction term that incorporates hesitation, proposing E_{AIFE} (Eq. (9)). Reconstructing the core variable into $\delta = (\mu - \nu)/(1 + \pi)$, E_{AIFE} enables hesitation to directly participate in the quantification of the entropy. For the above calculation example, $\delta_1 = 0.2/(1 + 0.2) \approx 0.1667$, $\delta_2 = 0.2/(1 + 0) = 0.2$. Because $|\delta_1| < |\delta_2|$, and AIFE is a monotonically decreasing function with respect to $|\delta|$ (see the proof in Section 3.1.3). So, with a higher hesitation, a higher entropy value is assigned. The design of AIFE enables hesitation to directly participate in entropy calculation, achieving a more refined and intuitive characterization of uncertainty. $E_{\mathrm{AIFE}}(a_1) > E_{\mathrm{AIFE}}(a_2)$. The results indicate that a_1 with higher hesitation is assigned a higher entropy value.

3.3.2. Mathematical definition

For an intuitionistic fuzzy set $A \in IFS_s(x)$, AIFE is defined as

$$E_{AIFE}(A) = \frac{1}{n}\sum_{i=1}^{n} \frac{\frac{4}{\pi}\arctan\left(\frac{1+\delta}{2}\right) + \frac{4}{\pi}\arctan\left(\frac{1-\delta}{2}\right) - 1}{\frac{8}{\pi}\arctan\left(\frac{1}{2}\right) - 1} \tag{9}$$

In which $\delta = \frac{\mu_A(x_i) - \nu_A(x_i)}{1 + \pi_A(x_i)}$. Especially, in the formula, π represents the ratio of the circumference of a circle to its diameter (3.14159…), while $\pi_A(x_i)$ denotes the uncertainty function in fuzzy sets.

3.3.3. Axiomatic proof

To verify that the proposed AIFE qualifies as a valid intuitionistic fuzzy entropy, it must be shown to satisfy the fundamental axioms of intuitionistic fuzzy entropy.

a. $A \in IFS_S(x)$, for $\forall x_i \in x = \{x_1, x_2, \cdots, x_n\}$, $E_{AIFE}(A) \subset [0, 1]$.

Proof. $\mu_A(x_i) \in [0,1], \nu_A(x_i) \in [0,1]$, $\pi_A(x_i) \in [0,1] \Rightarrow -1 \le \mu_A(x_i) - \nu_A(x_i) \le 1, 1 \le 1 + \pi_A(x_i) \le 2$, when $\mu_A(x_i) = 1, \nu_A(x_i) = 0$, the numerator is 1; when $\mu_A(x_i) = 0, \nu_A(x_i) = 1$, the numerator is -1. $|\delta| = \left|\frac{\mu_A - \nu_A}{1 + \pi_A}\right| \le \frac{1}{1} = 1$, therefore $|\delta| \le 1$, that is, $\delta \in [-1, 1]$.

Then let $f(\delta)=\frac{\frac{4}{\pi}\arctan\left(\frac{1+\delta}{2}\right)+\frac{4}{\pi}\arctan\left(\frac{1-\delta}{2}\right)-1}{\frac{8}{\pi}\arctan\left(\frac{1}{2}\right)-1}$, $D=\frac{8}{\pi}\arctan\left(\frac{1}{2}\right)-1$. Clearly, D is a constant, take the derivative with respect to δ, it can be calculated that $\frac{d}{d\delta}\arctan\left(\frac{1+\delta}{2}\right)=\frac{2}{4+(1+\delta)^2}$, $\frac{d}{d\delta}\arctan\left(\frac{1-\delta}{2}\right)=\frac{-2}{4+(1-\delta)^2}$, so $\frac{df(\delta)}{d\delta}=\frac{8}{\pi D}\left(\frac{1}{4+(1+\delta)^2}-\frac{1}{4+(1-\delta)^2}\right)$. For $\delta\in(0,1)$, $1+\delta>1-\delta>0\Rightarrow 4+(1+\delta)^2>4+(1-\delta)^2>0\Rightarrow\frac{1}{4+(1+\delta)^2}<\frac{1}{4+(1-\delta)^2}\Rightarrow\frac{1}{4+(1+\delta)^2}-\frac{1}{4+(1-\delta)^2}<0$. And $D>0$, thus $\frac{df(\delta)}{d\delta}<0$. Therefore, if $f(\delta)$ is monotonically decreasing with respect to δ in $[0,1]$, then $f(0)=\frac{1}{D}\left(\frac{4}{\pi}arctan\left(\frac{1}{2}\right)+\frac{4}{\pi}arctan\left(\frac{1}{2}\right)-1\right)=1$, $f(1)=\frac{1}{D}\left(\frac{4}{\pi}arctan(1)+\frac{4}{\pi}arctan(0)-1\right)=0$. Given the symmetry and that $f(\delta)$ is an even function, over the interval $[-1,0]$, $f(\delta)$ increases monotonically from $f(-1)=0$ to $f(0)=1$. Therefore, for $\delta\in[-1,1]$, $f(\delta)\subset[0,1]\Rightarrow E_{AIFE}(A)=\frac{1}{n}\sum_{i=1}^{n}f(\delta)$ represents the average value of numbers within the interval $[0,1]\Rightarrow E_{AIFE}(A)\subset[0,1]$.

b. $A\in IFS_s(X)$, for $\forall x_i\in X=\{x_1,x_2,\cdots,x_n\}$, the defined $E_{AIFE}(A)$ is an intuitionistic fuzzy entropy, it must satisfy the four conditions. □

Proof. The first condition. If A is a crisp set, that is, $A=\{\langle x,1,0\rangle|x\in A\}$ or $A=\{\langle x,0,1\rangle|x\in A\}$, then

$$E_{\mathrm{AIFE}}(A)=\frac{1}{n}\sum_{i=1}^{n}\frac{\frac{4}{\pi}\arctan\left(\frac{2}{2}\right)+\frac{4}{\pi}\arctan\left(\frac{0}{2}\right)-1}{\frac{8}{\pi}\arctan\left(\frac{1}{2}\right)-1}=0 \tag{10}$$

$$E_{\mathrm{AIFE}}(A)=\frac{1}{n}\sum_{i=1}^{n}\frac{\frac{4}{\pi}\arctan\left(\frac{0}{2}\right)+\frac{4}{\pi}\arctan\left(\frac{2}{2}\right)-1}{\frac{8}{\pi}\arctan\left(\frac{1}{2}\right)-1}=0 \tag{11}$$

In the following, we prove that if $E_{AIFE}(A)=0$, then A is a crisp set. Define the function $f(\delta)=\frac{\frac{4}{\pi}\arctan\left(\frac{1+\delta}{2}\right)+\frac{4}{\pi}\arctan\left(\frac{1-\delta}{2}\right)-1}{\frac{8}{\pi}\arctan\left(\frac{1}{2}\right)-1}$, then $E_{\mathrm{AIFE}}(A)=\frac{1}{n}\sum_{i=1}^{n}f(\delta)$. From the preceding proof, it follows that $f(\delta)\geq 0$ holds for $\delta\in[-1,1]$, and $f(\delta)=0$ holds if and only if $|\delta|=1$ holds, since $E_{\mathrm{AIFE}}(A)=0\Longrightarrow|\delta|=\left|\frac{\mu_A(x_i)-\nu_A(x_i)}{1+\pi_A(x_i)}\right|=1\Leftrightarrow|\mu_A(x_i)-\nu_A(x_i)|=1+\pi_A(x_i)$. Let $d_i=\mu_A(x_i)-\nu_A(x_i)$, $s_i=\mu_A(x_i)+\nu_A(x_i)\Rightarrow 1+\pi_A(x_i)=2-s_i,|d_i|=2-s_i$. Under intuitionistic fuzzy sets, $|\mu_A-\nu_A|\leq|\mu_A|+|\nu_A|=\mu_A+\nu_A\Rightarrow 0\leq s_i\leq 1$ and $|d_i|\leq s_i$, that is, $|d_i|=2-s_i\leq s_i\Rightarrow s_i\geq 1$. But $s_i\leq 1\Rightarrow s_i=1$, hence substituting yields $|d_i|=1$, namely $|\mu_A(x_i)-\nu_A(x_i)|=1$ and $\mu_A(x_i)+\nu_A(x_i)=1$. Solve the system of equations, if $\mu_A(x_i)-\nu_A(x_i)=1$ and $\mu_A(x_i)+\nu_A(x_i)=1\Rightarrow\mu_A(x_i)=1,\nu_A(x_i)=0$, if $\mu_A(x_i)-\nu_A(x_i)=-1$ and $\mu_A(x_i)+\nu_A(x_i)=1\Rightarrow\mu_A(x_i)=0,\nu_A(x_i)=1$. Then A is a crisp set.

The second condition. If $\mu_A(x_i)=\nu_A(x_i)\Rightarrow\delta=0\Rightarrow E_{\mathrm{AIFE}}(A)=\frac{1}{n}\sum_{i=1}^{n}\frac{\frac{4}{\pi}\arctan\left(\frac{1+0}{2}\right)+\frac{4}{\pi}\arctan\left(\frac{1-0}{2}\right)-1}{\frac{8}{\pi}\arctan\left(\frac{1}{2}\right)-1}=1$. In the following we prove that if $E_{\mathrm{AIFE}}(A)=1$, then by the preceding proof we know $E_{\mathrm{AIFE}}(A)=\frac{1}{n}\sum_{i=1}^{n}f(\delta)=1$, since $f(\delta)$ achieves its maximum value of 1 when $\delta=0$ holds, and $f(\delta)<1\Rightarrow$ holds for all $\delta\in[-1,1]$ when $\delta\neq 0$, while $f(\delta)=1$ holds if and only if $\delta=0$. Therefore, $\delta=0$, that is, for all $x_i\in X$, $\frac{\mu_A(x_i)-\nu_A(x_i)}{1+\pi_A(x_i)}=0\Rightarrow\mu_A(x_i)-\nu_A(x_i)=0\Longrightarrow\mu_A(x_i)=\nu_A(x_i)$.

The third condition. Let $A^c=\{\langle x,\nu_A(x),\mu_A(x)\rangle|x\in X\}$ denote the complement of A, then $\delta(A^c)=\frac{\nu_A(x_i)-\mu_A(x_i)}{1+\pi_A(x_i)}=-\delta(A)$. Therefore, $E_{\mathrm{AIFE}}(A^c)=\frac{1}{n}\sum_{i=1}^{n}\frac{\frac{4}{\pi}\arctan\left(\frac{1-\delta(A)}{2}\right)+\frac{4}{\pi}\arctan\left(\frac{1+\delta(A)}{2}\right)-1}{\frac{8}{\pi}\arctan\left(\frac{1}{2}\right)-1}=E_{\mathrm{AIFE}}(A)$.

The fourth condition. $\delta(A)=\frac{\mu_A(x_i)-\nu_A(x_i)}{1+\pi_A(x_i)}$, $\delta(B)=\frac{\mu_B(x_i)-\nu_B(x_i)}{1+\pi_B(x_i)}$, it has previously been demonstrated that $f(\delta)$ is a monotonically decreasing function of $|\delta|$, and $E_{\mathrm{AIFE}}(A)=\frac{1}{n}\sum_{i=1}^{n}f(\delta(A))$, $E_{\mathrm{AIFE}}(B)=\frac{1}{n}\sum_{i=1}^{n}f(\delta(B))$. When $\pi_A(x_i)=\pi_B(x_i)$, $f_A(x_i)\leq f_B(x_i)$, let $x_i\in X$ be fixed and let $\pi_A(x_i)=\pi_B(x_i)=\pi$. Then $|\delta_i(A)|=\left|\frac{\mu_A(x_i)-\nu_A(x_i)}{1+\pi}\right|=\frac{f_A(x_i)}{1+\pi}$. Given that $f_A(x_i)\leq f_B(x_i)$ implies $|\delta_i(A)|\leq\delta_i(B)|$, since $f(\delta)$ is a monotonically decreasing function of $|\delta|$, it follows that $E_{\mathrm{AIFE}}(A)=\frac{1}{n}\sum_{i=1}^{n}f(\delta_i(A))\geq\frac{1}{n}\sum_{i=1}^{n}f(\delta_i(B))=E_{\mathrm{AIFE}}(B)\Rightarrow E(A)\geq E(B)$.

Thus, the proposed AIFE entropy serves as an effective node splitting criterion in the subsequent AIFE-IIFDT, since it satisfies the axiomatic requirements for intuitionistic fuzzy entropy and provides a reliable measure of uncertainty in intuitionistic fuzzy sets. □

3.4. The FCM algorithm

An intuitionistic fuzzy decision tree cannot directly process raw data. To address this, the present study uses the algorithm [45], a classical fuzzy clustering method based on the optimization of the objective function, to transform raw data into a representation suitable for uncertainty modeling. The FCM algorithm effectively computes the membership of each sample to various cluster centers and has been widely adopted in the construction of fuzzy decision trees.

The fuzzification process is summarized in Fig. 2. For continuous features, FCM is applied to generate membership matrices. For discrete features, one-hot encoding is used to produce binary membership vectors. These steps provide the foundation for the subsequent calculation of the entropy and the construction of trees. It should be noted that only membership information is generated at this stage. Nonmembership and hesitation are derived in later steps through probabilistic conversion.

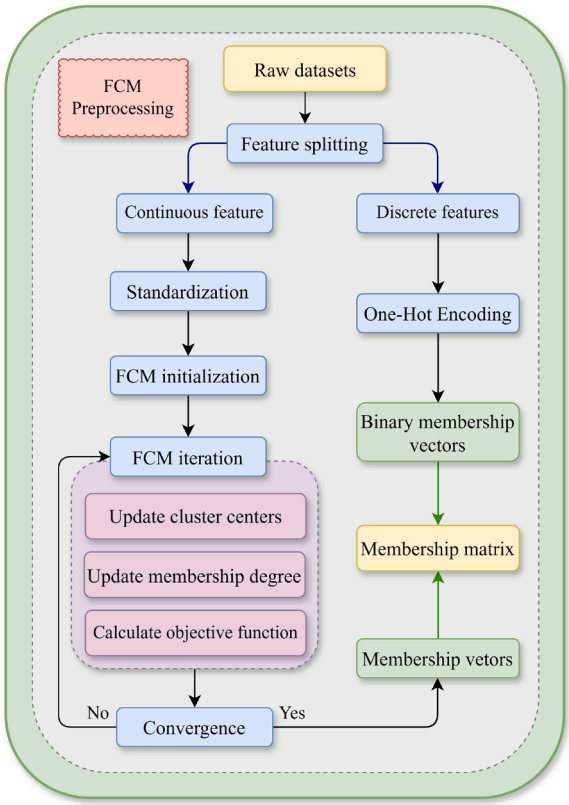


Fig. 2. FCM preprocessing.

Next is the parameter design for the FCM algorithm. The FCM algorithm achieves clustering by minimizing the following objective function.

$$J(U,V)=\sum_{i=1}^{n}\sum_{j=1}^{k}u_{ij}^{m}\parallel x_i-v_j\parallel^2 \tag{12}$$

where the parameters are defined as follows: n represents the total number of samples, k denotes the number of clusters, U is the membership matrix, V is the cluster center matrix, m denotes the fuzzy factor and $\parallel x_i-v_j\parallel^2$ represents the squared distance between sample x_i and cluster center v_j.

In this study, k is set to 2 for all cases, which means that each attribute is divided into two semantic categories, “low” and “high”. This provides clear semantics while reducing computational complexity. Setting $m=2$, the fuzzy factor controls the degree of fuzziness in the degree of membership. A value widely recommended in the literature [46] and validated through extensive experimentation to provide an effective balance between clustering performance and mathematical simplicity.

3.5. *Intuitionistic fuzzy decision tree with explicit hesitation*

This section mainly presents the core integration and methodological innovations of the AIFE-IIFDT model. The proposed model systematically combines the components introduced in the preceding sections. The fuzzified data obtained in Section 3.4 serve as input, while the AIFE proposed in Section 3.3 is used as a node splitting criterion. To enhance robustness, a threshold-based pruning mechanism using uncertainty measures is introduced to filter out highly uncertain rules, ultimately forming a complete and effective AIFE-IIFDT classification framework.

3.5.1. *Generate intuitionistic fuzzy representation*(μ,ν,π)

The FCM algorithm only generates membership. To construct an intuitionistic fuzzy decision tree, it is necessary to compute nonmembership and hesitation using the concepts of maximum support and maximum opposition. Let Pos^+ and Pos^- denote the maximum degrees of support and opposition. The conversion is performed as follows.

$$\begin{aligned} Pos^+&=\mu+\pi\\ Pos^-&=\nu+\pi \end{aligned} \tag{13}$$

Since IFS satisfy the condition that the sum of membership, nonmembership, and hesitation equals one, the following conversion can be derived.

$$\begin{aligned} \pi &= Pos^{+} + Pos^{-} - 1 \\ \mu &= Pos^{+} - \pi \\ \nu &= Pos^{-} - \pi \end{aligned} \tag{14}$$

The values of Pos^{+} and Pos^{-} are calculated on the basis of the mass distribution theorem proposed by Szmidt et al. [47], which converts probabilistic membership into intuitionistic fuzzy representations. This provides the necessary data foundation for subsequent intuitionistic fuzzy entropy calculation and intuitionistic fuzzy decision tree construction.

3.5.2. An overview of the basic framework of IIFDT

By integrating intuitionistic fuzzy set theory into the intuitionistic fuzzy decision tree framework, Li [35] developed the IIFDT, which not only accounts for membership and nonmembership but also explicitly incorporates hesitation, offering a more comprehensive characterization of uncertainty in decision-making. The IIFDT establishes a complete workflow spanning data preprocessing to rule extraction, making it suitable for multi-attribute decision-making and complex pattern recognition tasks.

The AIFE-IIFDT model proposed in this research is based on the IIFDT framework. The core innovation lies in replacing the original splitting criterion with the AIFE entropy proposed in this study, as detailed in Section 3.5.3.

3.5.3. The AIFE-IIFDT model construction

Following attribute fuzzification, we can build the AIFE-IIFDT model according to Algorithm 1. And the overall procedure of the proposed model is illustrated in Fig. 1b. The stopping criteria are as follows.

A node is marked as a leaf if any of the following conditions are met.

(1) Tree depth reaches maximum value. Splitting stops when the current node's depth exceeds the predefined maximum depth $depth_{\max}$.
(2) Insufficient node samples. Splitting is halted if the current node contains no samples.
(3) Node purity reached threshold. All samples within the current node belong to the same decision class (satisfying either $\mu \geq \beta_0$ or $\nu \geq \beta_0$), thus halting further splitting.
(4) Candidate feature set is empty. Splitting stops when no available features remain for splitting at the current node (all features have been utilized or are invalid). Thus splitting is halted.

And the pruning criteria are as follows.

(1) Significance level α (pre-pruning). Reduces noise introduced during data processing.
(2) The π threshold pruning (post-pruning). After a leaf node is generated, if its hesitation π exceeds the threshold, the node is discarded.

Algorithm 1 AIFE-IIFDT model construction. For input, fuzzified dataset D_f, significance level α, truth threshold β_0, hesitation threshold $\pi_{threshold}$, maximum tree depth $depth_{\max}$. For output, AIFE-IIFDT model.

1. Preprocessing. Perform preprocessing α on D_f, retaining only those with membership degree $\mu \geq \alpha$.
2. Conversion to an intuitionistic fuzzy representation. Compute the intuitionistic fuzzy representation (μ, ν, π) for each sample using the mass distribution function.
3. Recursive construction of decision trees.
  a. Calculate the AIFE values for candidate features and select the feature with the minimum entropy as the splitting node.
  b. If any stopping criterion is satisfied, the current node is a leaf.
4. Post-pruning. For leaf nodes, if $\pi \geq \pi_{threshold}$ or purity requirement β_0 is not satisfied, backtrack to the parent node for merging.
5. Rule extraction. Generate the rule set in the form, "Conditional Attribute Membership Constraint → Decision Class $C_k(\mu_k, \nu_k, \pi_k)$".
6. Model completion. Return the constructed AIFE-IIFDT model for classification and performance evaluation.

3.6. Example analysis of AIFE-IIFDT model

The preceding sections have outlined the construction of AIFE-IIFDT. The computational procedure for generating AIFE-IIFDT is demonstrated using datasets from the following small standard example library. The original fuzzy data set is shown in Table 2. Each condition attribute has been divided into two semantic subsets via FCM, and the decision attribute is binary classification. During calculations, the results are retained to four decimal places. For the denominator constant term of AIFE, do not directly calculate its value initially. If it can be simplified with the numerator, perform the simplification before calculating the specific value.

Table 2
Fuzzy data table.

Number	Attribute A		Attribute B		Decision Attribute C	
	A_1	A_2	B_1	B_2	C^+	C^-
1	1.00	0.00	1.00	0.00	0.00	1.00
2	1.00	0.00	0.00	1.00	1.00	0.00
3	0.59	0.41	1.00	0.00	0.00	1.00
4	0.01	0.99	0.88	0.12	1.00	0.00

3.6.1. Calculate intuitionistic fuzzy entropy of attribute A

The attribute A has values A_1 and A_2, calculate relative frequency $p(C|A_i)$. $w(C^+ \wedge A_1) = 1.01, w(C^+ \wedge A_2) = 0.99, w(C^- \wedge A_1) = 1.59, w(C^- \wedge A_2) = 0.41$, $p(C^+ \mid A_1) = \frac{w(C^+ \wedge A_1)}{w(C^+ \wedge A)} = \frac{1.01}{2} = 0.505, p(C^- \mid A_1) = \frac{w(C^- \wedge A_1)}{w(C^- \wedge A)} = \frac{1.59}{2} = 0.795$, $p(C^+ \mid A_2) = \frac{w(C^+ \wedge A_2)}{w(C^+ \wedge A)} = \frac{0.99}{2} = 0.495, p(C^- \mid A_2) = \frac{w(C^- \wedge A_2)}{w(C^- \wedge A)} = \frac{0.41}{2} = 0.205$.

Calculate the maximum support Pos^+ and maximum opposition Pos^- of the attribute A using Szmidt's quality distribution theorem. For category C^+, the sorting $p(C^+ \mid A_i)$ in descending order yields $p_1 = 0.505(A_1), p_2 = 0.495(A_2)(0 \le p_{i+1} \le p_i \le 1, i = 1, 2)$, and $\sum_{i=1}^{2} p_i = 1$. $m(G_2) = \mu_2, G_2 = \{x \in \Omega | P(x) \ge p_2 = 0.495\}, m(G_1) = \mu_1 - \mu_2, G_1 = \{x \in \Omega | P(x) \ge p_1 = 0.505\}$. Calculate μ_i using formula $\mu_i = |G_i| p_i + \sum_{j=i+1}^{n} \left(|G_j| - |G_{j-1}| \right) p_j$, $\mu_2 = 2p_2 = 0.99, \mu_1 = p_1 + p_2 = 1$. For category C^-, the sorting $p(C^- \mid A_i)$ in descending order yields $p_1 = 0.795(A_1), p_2 = 0.205(A_2)$. Similarly, μ_i is calculated using formula. $\mu_2 = 2 \times 0.205 = 0.41$, $\mu_1 = 0.795 + 0.205 = 1$. Therefore, $Pos^-(A_1) = 1, Pos^-(A_2) = 0.41$.

Calculate (μ, ν, π) corresponding to A_1 and A_2. For A_1, $\pi(A_1) = Pos^+(A_1) + Pos^-(A_1) - 1 = 1, \mu(A_1) = Pos^+(A_1) - \pi(A_1) = 0, \nu(A_1) = Pos^-(A_1) - \pi(A_1) = 0$. For A_2, $\pi(A_2) = Pos^+(A_2) + Pos^-(A_2) - 1 = 0.4, \mu(A_2) = Pos^+(A_2) - \pi(A_2) = 0.59, \nu(A_2) = Pos^-(A_2) - \pi(A_2) = 0.01$.

Calculate $E_{AIFE}(A)$ according to Formula (9). $E_{\text{AIFE}}(A_1) = 1, E_{\text{AIFE}}(A_2) \approx 0.8106 \Rightarrow E_{\text{AIFE}}(A) = \frac{1}{2}\left(E_{\text{AIFE}}(A_1) + E_{\text{AIFE}}(A_2)\right) \approx 0.9053$.

3.6.2. Calculate intuitionistic fuzzy entropy of attribute B

All steps are consistent with the intuitionistic fuzzy entropy of attribute A, so the formulas and calculation processes will not be detailed hereafter.

The attribute B has values B_1 and B_2, calculate relative frequency $p(C|B_i)$. $w(C^+ \wedge B_1) = 0.88, w(C^+ \wedge B_2) = 1.12, w(C^- \wedge B_1) = 2, w(C^- \wedge B_2) = 0$, $p(C^+ \mid B_1) = 0.44, p(C^- \mid B_1) = 1.00, p(C^+ \mid B_2) = 0.56, p(C^- \mid B_2) = 0.00$.

Calculate Pos^+ and Pos^- of the attribute B using Szmidt's quality distribution theorem. For category C^+, descending order yields $p_1 = 1.00(B_1)$, $p_2 = 0.00(B_2), \mu_2 = 0, \mu_1 = 1$. So $Pos^-(B_1) = 1, Pos^-(B_2) = 0$.

Calculate (μ, ν, π) corresponding to B_1 and B_2. For B_1, $\pi = 0.88, \mu = 0, \nu = 0.12$. For B_2, $\pi = 0, \mu = 1, \nu = 0$.

Calculate $E_{AIFE}(B)$ according to Formula (9). $E_{\text{AIFE}}(B_1) \approx 0.9950, E_{\text{AIFE}}(B_2) = 0 \Rightarrow E_{\text{AIFE}}(B) = \frac{1}{2}\left(E_{\text{AIFE}}(B_1) + E_{\text{AIFE}}(B_2)\right) = 0.4975$.

3.6.3. Root node selection and node pruning

Comparing the AIFE of two attributes, since $E_{\text{AIFE}}(A) = 0.9053 > E_{\text{AIFE}}(B) = 0.4975$, attribute B has lower intuitionistic fuzzy entropy, indicating lower uncertainty. Therefore, attribute B is selected as the extended attribute (root node). Assuming $\beta_0 = 0.75$, the preceding calculations yielded the μ and ν of B_1 and B_2. The μ and ν of B_1 are 0 and 0.12, both less than 0. Therefore, further partitioning of this node is required. The μ and ν of B_2 are 1 and 0. Since the μ is greater than β_0, splitting is terminated and it is marked as a leaf node, the classification decision is C^+.

3.6.4. Calculate intuitionistic fuzzy entropy of attribute A under the B_1 condition

Next, calculate the fuzzy entropy of the fuzzy partitions A_1 and A_2 for attribute A under condition attribute B_1. Proceed with the calculation step by step as described above.

Calculate the relative frequency of A_1 and A_1 relative to B_1. $w(C^+ \wedge B_1 \wedge A_1) = 0.0088, w(C^+ \wedge B_1 \wedge A_2) \approx 0.8712, w(C^+ \wedge B_1) = 0.88, w(C^- \wedge B_1 \wedge A_1) = 1.59, w(C^- \wedge B_1 \wedge A_2) = 0.41, w(C^- \wedge B_1) = 2.00$. $p(C^+ \mid B_1 \wedge A_1) = \frac{w(C^+ \wedge B_1 \wedge A_1)}{w(C^+ \wedge B_1 \wedge A)} = 0.01, p(C^- \mid B_1 \wedge A_1) = \frac{w(C^- \wedge B_1 \wedge A_1)}{w(C^- \wedge B_1 \wedge A)} = 0.795, p(C^+ \mid B_1 \wedge A_2) = \frac{w(C^+ \wedge B_1 \wedge A_2)}{w(C^+ \wedge B_1 \wedge A)} = 0.99, p(C^- \mid B_1 \wedge A_2) = \frac{w(C^- \wedge B_1 \wedge A_2)}{w(C^- \wedge B_1 \wedge A)} = 0.205$.

Calculate Pos^+ and Pos^-. For category C^+, descending order yields $p_1 = 0.99(A_2), p_2 = 0.01(A_1), \mu_2 = 0.02, \mu_1 = 1$. So $Pos^+(B_1 \wedge A_1) = 0.02, Pos^+(B_1 \wedge A_2) = 1$. For category C^-, descending order yields $p_1 = 0.795(A_1), p_2 = 0.205(A_2), \mu_2 = 0.41, \mu_1 = 1$. $Pos^-(B_1 \wedge A_1) = 1, Pos^-(B_1 \wedge A_2) = 0.41$.

Calculate the (μ, ν, π). For $B_1 \wedge A_1$, $\pi = 0.02, \mu = 0, \nu = 0.98$. For $B_1 \wedge A_2$, $\pi = 0.41, \mu = 0.59, \nu = 0$.

Calculate $E_{\text{AIFE}}(B_1 \wedge A)$ according to Formula (9). $E_{\text{AIFE}}(B_1 \wedge A_1) \approx 0.0682, E_{\text{AIFE}}(B_1 \wedge A_2) \approx 0.8070 \Rightarrow E_{\text{AIFE}}(B_1 \wedge A) \approx 0.4376$.

3.6.5. Final AIFE-IIFDT model

Since all attributes have been exhausted, the leaf nodes can be directly determined. The conditional attribute is B_1, when attribute A is A_1, it supports C^-; when attribute A is A_2, it supports C^+. Therefore, the final decision tree is formed as shown in Fig. 3. The path from the root node to each leaf node in the diagram constitutes a classification rule, resulting in three rules.

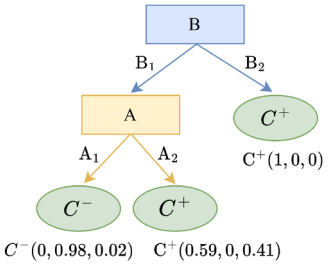


Fig. 3. AIFE-IIFDT model constructed using example data.

(1) Rule 1. If B is B_2, then classify it as C^+.
(2) Rule 2. If B is B_1 and A is A_1, then classify it as C^-.
(3) Rule 3. If B is B_1 and A is A_2, then classify it as C^+.

3.7. Time complexity analysis of the AIFE-IIFDT model

This study provides a theoretical analysis of the computational time complexity of the proposed AIFE-IIFDT model. The overall complexity of the algorithm consists of two main components. First, data fuzzification based on the FCM algorithm. Second, the construction of AIFE-IIFDT model.

(1) Data Fuzzification.

Let N denote the total number of samples and D denote the total number of original features. Where D_c is the number of continuous features and D_s is the number of categorical features, satisfying $D = D_c + D_s$. FCM is applied only to D_c continuous features, dividing each feature into K fuzzy clusters (where $K = 2$ in this work). Let T_f denote the maximum number of iterations for FCM.

When performing FCM on a single continuous feature, the core operations in each iteration include, calculating the distances ($O(N \cdot K)$) from N samples to K cluster centers, and updating the $N \times K$ dimensional membership matrix ($O(N \cdot K^2)$). Therefore, the complexity of processing a single feature is $O(T_f \cdot N \cdot K^2)$. Since the clustering process for each continuous feature is independent, the overall time complexity for this step is $O_{FCM} = O(T_f \cdot D_c \cdot N \cdot K^2)$. Categorical features undergo one-hot encoding processing with complexity $O(D_s \cdot N \cdot L)$, where L represents the average number of categories for categorical features. This complexity is a minor factor in the overall analysis and can be considered negligible relative to the FCM process.

(2) Construction of AIFE-IIFDT Model.

After fuzzification, the feature dimension used to construct the decision tree expands to $F = D_c \cdot K + D_s \cdot L$. Decision trees perform node splits by recursively selecting the feature with the smallest AIFE entropy. Let the maximum tree depth be H.

For each internal node, the algorithm must evaluate all F candidate features. For a feature, calculating its AIFE requires sequentially performing the following steps, calculate support, determine maximum likelihood, derive the fuzzy intuitionistic expression (μ, ν, π), and finally calculate the AIFE. Therefore, the complexity of selecting the optimal split point on a node containing N' samples is $O(F \cdot N')$. Under the assumption that the decision tree is nearly balanced, the total complexity of constructing the entire tree can be approximated as $O_{\text{Tree}} = O(H \cdot F \cdot N)$.

(3) Total Time Complexity.

Combining the above two steps, the overall time complexity of the AIFE-IIFDT model is $O_{\text{Total}} = O_{\text{FCM}} + O_{\text{Tree}} = O(T_f \cdot D_c \cdot N \cdot K^2 + H \cdot F \cdot N)$. Substituting $F = D_c \cdot K + D_s \cdot L$, yields $O_{\text{Total}} = O\big(N \cdot D_c \cdot (T_f \cdot K^2 + H \cdot K) + N \cdot H \cdot D_s \cdot L\big)$.

Analysis indicates that the computational cost of the algorithm exhibits a linear relationship with both the sample size N and the number of features D_c and D_s. Among these, the FCM clustering iterations T_f, the number of clusters K, the tree depth H, and the average number of categories L are all predefined constants. Therefore, AIFE-IIFDT model exhibits excellent scalability.

4. Experimental studies

This chapter conducts systematic experimental validation and performance evaluation of the proposed AIFE-IIFDT model. Section 4.1 details the experimental setup, including evaluation metrics, parameter settings, and dataset descriptions. Section 4.2 presents the experimental results from multiple perspectives, followed by analysis and discussion.

4.1. Experimental setup

This section systematically introduces the experimental setup to comprehensively evaluate the classification performance of the AIFE-IIFDT model and its effectiveness in handling uncertain information. To ensure the reliability of statistical results, all comparative experiments were conducted with 30 independent repetitions, reporting both the mean and standard deviation. The experimental environment consists of: Intel Core i7-1065G7 CPU @ 1.30GHz, 1.50GHz, 16.0 GB RAM, NVIDIA GeForce MX350 graphics card, running on a 64-bit Windows operating system. All algorithms were implemented using Python within the Anaconda environment.

Table 3
Parameter design for six intuitionistic fuzzy entropy methods.

Parameter	Value
Number of clusters k	$k = 2$
Fuzziness factor m	$m = 2$
Significance level α	$\alpha \in [0, 0.5]$
Authenticity threshold β_0	$\beta_0 \in [0.5, 1]$
Hesitation threshold π	$\pi = 0.1$
Maximum tree depth $D_{\max}$	$D_{\max} = 7$

Table 4
Parameter settings for seven comparison algorithms.

Algorithms	Parameter
C4.5	$D_{\max} = 7$
CART	$D_{\max} = 7$
RBFNN	hidden_neurons = 50, n_epochs = 300, $\eta = 0.01$
RF	$D_{\max} = 7$, n_estimators = 100, subsample = 0.8
GBDT	$D_{\max} = 7$, n_estimators = 100, $\eta = 0.1$
XGBoost	$D_{\max} = 7$, $\eta = 0.1$, num_boost_round = 100
IFDT	$C \in [2, 3, 4, 5, 6, 7]$, $S \in [2, 3, 5, 7]$, $d_\pi \in [0.5, 0.7, 0.9]$, $n_\alpha = 5$, $d_\beta \in [5, 7]$

4.1.1. Evaluation metrics

This study selected five widely used evaluation metrics: Accuracy, Precision, Recall, F1-Score, and Area Under the Curve (AUC). The definitions are as follows.

Among them, true positives (TP) denotes the number of actually positive samples that are correctly predicted as positive, true negatives (TN) refer to the number of actually negative samples that are correctly predicted as negative, false positives (FP) and false negatives (FN) represent the numbers of misclassified positive and negative samples, respectively.

(1) Accuracy. The proportion of correctly classified samples relative to the total number of samples.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

(2) Precision. The proportion of samples predicted as positive that are actually positive.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{16}$$

(3) Recall. The proportion of samples that are actually positive and that are correctly predicted.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{17}$$

(4) F1-Score. The harmonic mean of precision and recall.

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{18}$$

(5) AUC. Area under the receiver operating characteristic curve, reflecting the model's overall ability to distinguish between positive and negative classes across different thresholds. A value closer to 1 indicates better performance.

4.1.2. Parameter settings

For the six intuitionistic fuzzy entropy methods embedded into decision trees, grid search and five-fold cross-validation were employed to optimize key parameters with the objective of maximizing AUC. Grid search within the preset parameter space determined the optimal parameter combinations. The parameter settings for the six entropy methods are detailed in Table 3, while the parameter settings for the seven comparative classification algorithms are detailed in Table 4. To balance the Accuracy of parameter optimization with the fairness of time comparisons among algorithms, the grid search step size was set to 0.01 when evaluating classification performance to obtain finely tuned optimal parameters. When measuring runtime, the step size was set to 0.1 to ensure the parameter adjustment granularity of the proposed method remained comparable to other algorithms, thereby avoiding unreasonable time overhead caused by overly refined searches.

4.1.3. Public dataset

This research selected 12 binary classification datasets from the UCI and KEEL databases for experimental validation. The basic information of the selected datasets is summarized in Table 5. The column of feature composition in the table uses R(real), I(integer) and C(categorical) to represent the distribution of attributes of each dataset.The choice of a binary classification task is primarily based on two reasons: the calculation of intuitionistic fuzzy entropy relies on explicit semantic associations between membership,

Table 5
Introduction to the dataset.

Dataset	Samples	Feature(R,I,C)
Gallstone (GAL)	319	38(28,3,7)
Breast Cancer Wisconsin Prognostic (BCW)	198	33(30,3,0)
Ionosphere (ION)	351	34(33,1,0)
Breast Cancer Wisconsin Diagnostic (WDB)	569	30(30,0,0)
Pima Indians Diabetes data set (PID)	768	8(8,0,0)
Appendicitis (APP)	106	7(7,0,0)
South African Heart data set (SAH)	462	9(5,3,1)
Heart Failure Clinical Records (HFC)	299	12(2,5,5)
Hepatitis (HEP)	155	19(2,17,0)
Heart Cleveland Disease (HCD)	303	13(1,7,5)
Glioma Grading Clinical and Mutation Features (GGC)	839	23(1,0,22)
Indian Liver Patient Dataset (ILP)	538	10(5,4,1)

nonmembership, and hesitation. In binary classification, these three elements directly correspond to the positive class, negative class, and uncertainty, resulting in a clear structure. Additionally, binary classification problems exhibit relatively low computational complexity, making them suitable as benchmarks for evaluating algorithm performance.

4.2. Experimental results

This section analyzes and discusses the experimental results. First, the proposed AIFE entropy is compared with five existing intuitionistic fuzzy entropies within the same decision tree framework to validate the effectiveness of AIFE as a splitting criterion. Subsequently, AIFE-IIFDT was comprehensively compared with seven representative classification algorithms across multiple public datasets. Finally, the model performance is further evaluated across dimensions including statistical significance, feature interpretability, computational efficiency, and scalability.

4.2.1. Comparative experiment of six different entropies

To validate the effectiveness of AIFE in intuitionistic fuzzy decision trees, this study conducted a comparative analysis with five recently proposed intuitionistic fuzzy entropies, including E_{LY} (2019), E_{XL} (2023), E_{JQ} (2023), E_{KS} (2024) and E_{LS} (2025). These five types of entropy were initially applied in different scenarios. This study incorporates them as splitting criteria into decision trees to examine their applicability in classification tasks. Table 6 presents the mean and standard deviation of each method in five metrics: Accuracy, Precision, Recall, F1-Score, and AUC (with optimal values highlighted in bold).

As shown in the experimental results presented in Table 6, different entropy functions exhibit varying performance across datasets. Overall, the proposed AIFE demonstrates outstanding performance across multiple datasets and evaluation metrics. Taking Accuracy as an example, AIFE achieved optimal values of 0.8427, 0.9121, 0.9067, 0.9503, 0.9979, and 0.9372 on the GAL, APP, HFC, WDB, HEP, and HCD datasets. In addition, AIFE achieved the optimal values on seven, six, and five datasets respectively for the Precision, F1-Score and AUC. In terms of standard deviation, AIFE entropy exhibits acceptable variability across most datasets. The remaining five entropy measures demonstrate varying strengths across multiple datasets, indicating that different entropy functions adapt differently to specific data characteristics.

Based on a single representative experiment, Fig. 4 visually illustrates the distribution of five metrics across ten datasets to compare the performance profiles of different methods. Fig. 5 presents the corresponding ROC curves for each dataset, enabling an intuitive assessment of overall discrimination capability. In summary, incorporating the intuitionistic fuzzy entropy from multi-attribute decision-making and other scenarios into decision tree splitting criteria is feasible, with different entropy functions exhibiting distinct advantages under varying data distributions. The proposed AIFE entropy achieves optimal performance across most datasets by explicitly integrating hesitation, validating its effectiveness and universality as a splitting criterion.

4.2.2. Comparative experiment of eight different algorithms

This section comprehensively compares the proposed AIFE-IIFDT model with seven representative classification algorithms across 12 datasets, evaluating their performance across three dimensions: classification accuracy, statistical significance, and computational overhead. The selected comparison algorithms encompass diverse modeling methods. C4.5 [39] and CART [40], as classic deterministic decision trees, employ information gain and Gini-index as splitting criterion respectively. RF [41], GBDT [42], and XGBoost [43] are powerful ensemble learning algorithms. RBFNN [44] represents a fundamental neural network model. Additionally, the IFDT proposed by Ren et al. [48] is included in the comparison as a representative method from recent years. This comprehensive comparison aims to thoroughly evaluate the generalization capability and robustness of the proposed model under different data structures.

As shown in Table 7, the mean and standard deviation results for Accuracy indicate that AIFE-IIFDT achieved optimal means of 0.8427, 0.9121, 0.8883, 0.9067, 0.9979, 0.9331, 0.9372, and 0.8724. In the remaining four datasets, RBFNN achieved the highest accuracy on SAH and ILP, while XGBoost outperformed on ION and WDB. In terms of standard deviation, AIFE-IIFDT demonstrates minimal fluctuation that is either superior to or comparable with the benchmark methods across multiple datasets. For example,

Table 6
Comparison results for six different entropy types.

Entropy	GAL									
	Accuracy	Std	Precision	Std	Recall	Std	F1-Score	Std	AUC	Std
E_{LY}	0.7052	0.0547	0.7003	0.0540	0.7250	0.1246	0.7066	0.0688	0.8075	0.0637
E_{JQ}	0.7943	0.0520	0.7747	0.0527	0.8333	0.0667	0.8017	0.0504	0.8427	0.0332
E_{KS}	0.8302	0.0465	0.8027	0.0474	0.8802	0.0646	0.8381	0.0433	0.8771	0.0414
E_{LS}	0.7781	0.0537	0.7695	0.0574	0.8000	0.0702	0.7827	0.0505	0.8476	0.0420
E_{XL}	0.8266	**0.0428**	0.7927	0.0463	**0.8875**	**0.0446**	0.8368	**0.0385**	0.8641	0.0439
$\mathbf{E}_{\mathbf{AIFE}}$	**0.8427**	0.0456	**0.8343**	**0.0322**	0.8563	0.0911	**0.8426**	0.0526	**0.8878**	**0.0315**
Entropy	APP									
	Accuracy	Std	Precision	Std	Recall	Std	F1-Score	Std	AUC	Std
E_{LY}	0.8803	0.0491	0.8885	0.0388	0.9796	0.0564	0.9302	0.0307	0.8273	0.1444
E_{JQ}	0.8894	0.0494	0.8934	0.0415	0.9852	0.0429	0.9359	0.0292	0.8681	0.1443
E_{KS}	0.9045	0.0503	0.9057	0.0443	0.9889	0.0222	0.9448	0.0279	**0.8792**	0.1437
E_{LS}	0.8803	0.0491	0.8915	0.0412	0.9759	0.0568	0.9300	0.0306	0.8664	0.1474
E_{XL}	0.8864	**0.0435**	0.8888	0.0399	0.9870	0.0235	0.9347	**0.0240**	0.8711	0.1461
$\mathbf{E}_{\mathbf{AIFE}}$	**0.9121**	0.0524	**0.9059**	0.0506	**1.0000**	**0.0000**	**0.9499**	0.0283	0.8451	**0.1298**
Entropy	BCW									
	Accuracy	Std	Precision	Std	Recall	Std	F1-Score	Std	AUC	Std
E_{LY}	0.9108	0.0670	0.9157	0.0615	0.9796	0.0412	0.9452	0.0407	0.9474	0.0636
E_{JQ}	0.8825	0.0626	0.9148	0.0464	0.9366	0.0395	0.9253	0.0401	0.9454	0.0513
E_{KS}	0.9483	0.0581	0.9551	0.0562	0.9828	0.0273	0.9679	0.0357	0.9693	0.0392
E_{LS}	0.9233	0.0854	0.9452	0.0663	0.9591	0.0471	0.9516	0.0535	0.9474	0.0665
E_{XL}	**0.9583**	**0.0522**	**0.9591**	0.0518	**0.9914**	**0.0235**	**0.9743**	**0.0321**	**0.9806**	**0.0342**
$\mathbf{E}_{\mathbf{AIFE}}$	0.8883	0.0527	0.9266	**0.0378**	0.9301	0.0382	0.9281	0.0342	0.9450	0.0391
Entropy	SAH									
	Accuracy	fStd	Precision	Std	Recall	Std	F1-Score	Std	AUC	Std
E_{LY}	0.8530	0.0398	**0.9183**	0.0257	0.8525	0.0591	0.8829	0.0350	0.9419	0.0254
E_{JQ}	**0.8695**	0.0308	0.9022	0.0132	**0.8984**	0.0424	**0.8998**	0.0266	**0.9484**	0.0230
E_{KS}	0.8530	**0.0134**	0.8831	0.0111	0.8945	**0.0183**	0.8886	**0.0106**	0.9311	0.0148
E_{LS}	0.8391	0.0313	0.9140	**0.0106**	0.8333	0.0549	0.8707	0.0295	0.9430	**0.0097**
E_{XL}	0.8534	0.0143	0.8836	0.0117	0.8945	0.0185	0.8889	0.0112	0.9311	0.0148
$\mathbf{E}_{\mathbf{AIFE}}$	0.8513	0.0182	0.8847	0.0125	0.8896	0.0344	0.8866	0.0162	0.9346	0.0179
Entropy	ION									
	Accuracy	Std	Precision	Std	Recall	Std	F1-Score	Std	AUC	Std
E_{LY}	0.9174	0.0414	0.8751	0.0891	0.9080	0.0595	0.8872	0.0520	0.9588	0.0270
E_{JQ}	0.9479	0.0324	0.9271	0.0772	0.9333	**0.0189**	0.9283	0.0404	0.9723	0.0110
E_{KS}	0.9671	0.0190	0.9574	0.0489	**0.9520**	0.0299	0.9537	0.0258	0.9925	0.0102
E_{LS}	0.9225	0.0339	0.8584	0.0766	0.9453	0.0193	0.8977	0.0404	0.9725	0.0089
E_{XL}	**0.9685**	**0.0144**	**0.9619**	**0.0417**	0.9507	0.0246	**0.9554**	**0.0196**	**0.9937**	**0.0066**
$\mathbf{E}_{\mathbf{AIFE}}$	0.9113	0.0457	0.8391	0.1008	0.9467	0.0316	0.8858	0.0545	0.9690	0.0108
Entropy	HFC									
	Accuracy	Std	Precision	Std	Recall	Std	F1-Score	Std	AUC	Std
E_{LY}	0.8983	**0.0386**	0.8831	0.0494	0.9862	0.0224	0.9306	**0.0240**	0.9859	**0.0100**
E_{JQ}	0.9039	0.0423	0.8797	0.0464	**0.9992**	**0.0044**	0.9350	0.0269	**0.9880**	0.0166
E_{KS}	0.9022	0.0412	0.8786	0.0451	0.9976	0.0099	0.9337	0.0267	0.9740	0.0230
E_{LS}	0.8922	0.0483	0.8766	0.0565	0.9870	0.0294	0.9270	0.0300	0.9833	0.0157
E_{XL}	0.9017	0.0411	0.8803	**0.0421**	0.9935	0.0177	0.9330	0.0272	0.9737	0.0235
$\mathbf{E}_{\mathbf{AIFE}}$	**0.9067**	0.0453	**0.8861**	0.0505	0.9951	0.0098	**0.9366**	0.0289	0.9842	0.0191
Entropy	WDB									
	Accuracy	Std	Precision	Std	Recall	Std	F1-Score	Std	AUC	Std
E_{LY}	0.9351	0.0080	0.9126	**0.0125**	**0.9926**	0.0086	0.9508	0.0059	0.9778	**0.0077**
E_{JQ}	0.9351	0.0125	0.9215	0.0138	0.9810	0.0116	0.9502	0.0095	0.9704	0.0145
E_{KS}	0.9351	0.0108	0.9450	0.0336	0.9551	0.0308	0.9490	0.0079	0.9706	0.0154
E_{LS}	0.9348	0.0100	0.9193	0.0139	0.9833	**0.0083**	0.9502	0.0074	0.9669	0.0175
E_{XL}	0.9380	0.0068	0.9224	0.0144	0.9852	0.0148	0.9526	0.0051	0.9698	0.0146
$\mathbf{E}_{\mathbf{AIFE}}$	**0.9503**	**0.0047**	**0.9806**	0.0139	0.9403	0.0130	**0.9598**	**0.0036**	**0.9791**	0.0107

Table 6
continued.

Entropy	HEP									
	Accuracy	Std	Precision	Std	Recall	Std	F1-Score	Std	AUC	Std
E_{LY}	0.9292	0.0212	0.7300	0.0510	**1.0000**	**0.0000**	0.8429	0.0364	0.9564	**0.0131**
E_{JQ}	0.9333	0.0509	0.7683	0.1562	0.9889	0.0598	0.8571	0.1059	0.9675	0.0444
E_{KS}	0.9438	0.0493	0.8017	0.1620	**1.0000**	**0.0000**	0.8810	0.0999	0.9769	0.0274
E_{LS}	0.9354	0.0411	0.7650	0.1324	**1.0000**	**0.0000**	0.8607	0.0821	0.9739	0.0260
E_{XL}	0.9188	0.0649	0.7393	0.1779	0.9778	0.0831	0.8319	0.1276	0.9585	0.0579
$\mathbf{E}_{\mathbf{AIFE}}$	**0.9979**	**0.0112**	**1.0000**	**0.0000**	0.9889	0.0598	**0.9933**	**0.0359**	**0.9944**	0.0299
Entropy	PID									
	Accuracy	Std	Precision	Std	Recall	Std	F1-Score	Std	AUC	Std
E_{LY}	0.9236	0.0274	0.9612	**0.0097**	0.9193	0.0373	0.9395	0.0235	0.9853	0.0083
E_{JQ}	**0.9604**	0.0212	0.9627	0.0168	0.9773	0.0349	**0.9695**	0.0169	0.9852	**0.0022**
E_{KS}	0.9539	0.0199	0.9445	0.0113	0.9870	0.0261	0.9651	0.0155	0.9808	0.0065
E_{LS}	0.9543	**0.0181**	0.9429	0.0120	**0.9897**	0.0237	0.9656	**0.0141**	0.9820	0.0062
E_{XL}	0.9541	0.0200	0.9448	0.0113	0.9870	0.0261	0.9653	0.0156	0.9808	0.0067
$\mathbf{E}_{\mathbf{AIFE}}$	0.9331	0.0200	**0.9735**	0.0267	0.9230	**0.0230**	0.9472	0.0155	**0.9876**	0.0100
Entropy	HCD									
	Accuracy	Std	Precision	Std	Recall	Std	F1-Score	Std	AUC	Std
E_{LY}	0.8503	0.0162	0.8750	0.0237	0.8444	0.0103	0.8593	0.0140	0.9317	0.0137
E_{JQ}	0.8530	0.0192	0.8782	0.0251	0.8465	0.0259	0.8617	0.0176	0.9334	0.0172
E_{KS}	0.8639	0.0208	0.8823	0.0207	0.8646	0.0412	0.8727	0.0205	0.9396	0.0121
E_{LS}	0.8525	**0.0120**	0.8762	0.0194	0.8475	0.0054	0.8615	**0.0099**	0.9356	**0.0093**
E_{XL}	0.8645	0.0211	0.8830	**0.0188**	0.8646	0.0412	0.8731	0.0209	0.9405	0.0106
$\mathbf{E}_{\mathbf{AIFE}}$	**0.9372**	0.0170	**0.8986**	0.0218	**0.9970**	**0.0091**	**0.9451**	0.0142	**0.9697**	0.0106

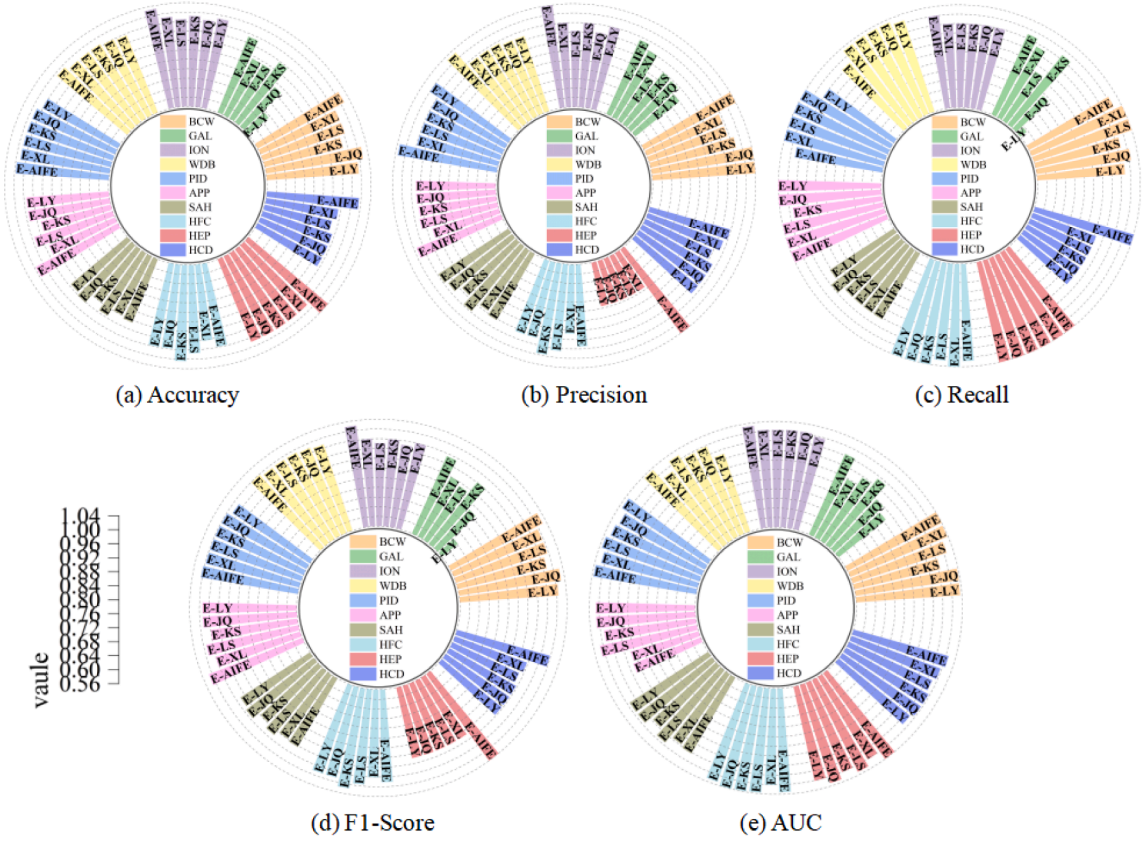


Fig. 4. Visualization of five evaluation metrics for ten datasets across six different methods. (a) Accuracy. (b) Precision. (c) Recall. (d) F1-Score. (e) AUC.

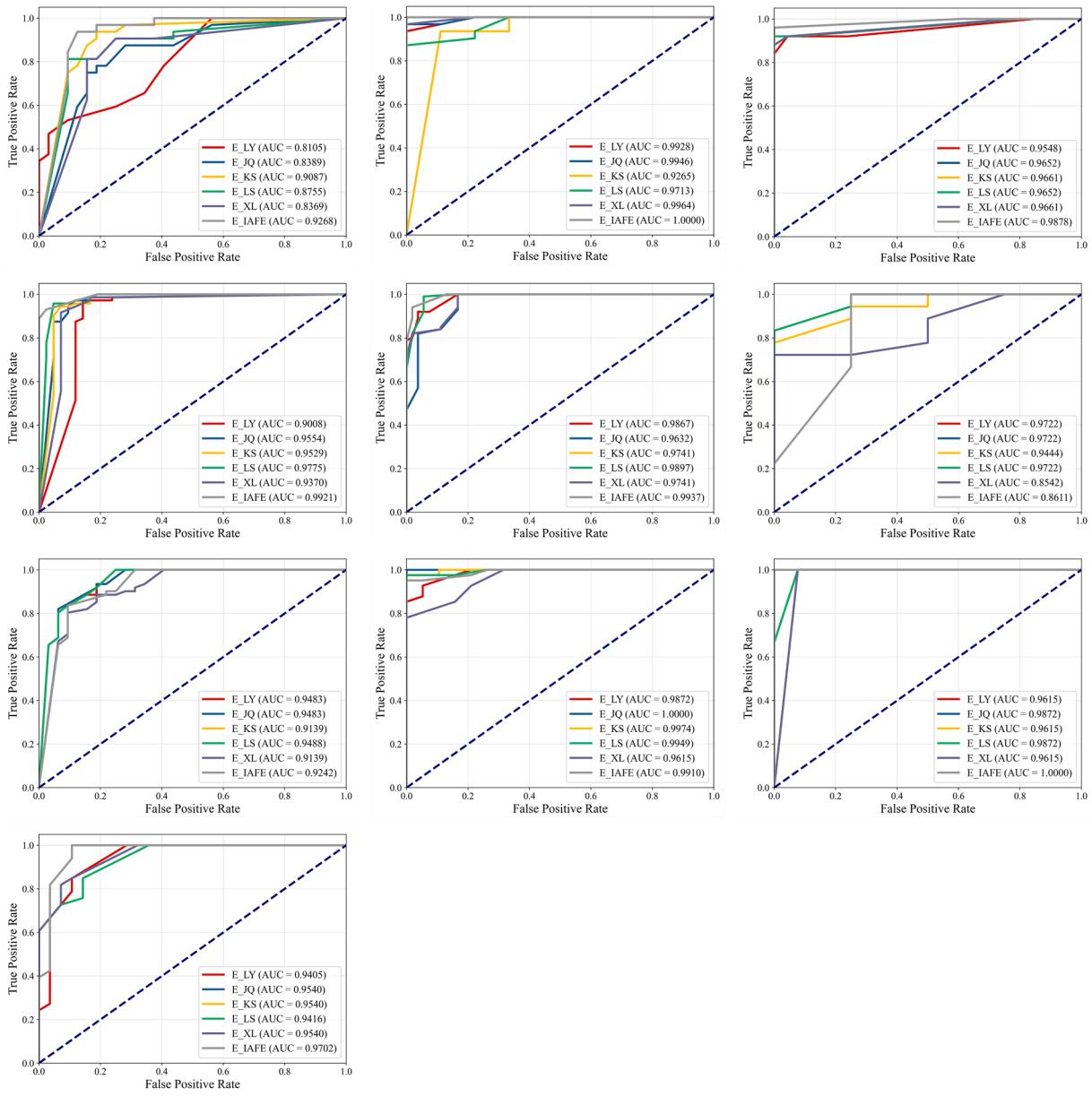


Fig. 5. The ROC curves for ten datasets, demonstrating the classification performance in distinguishing positive and negative samples. The pictures are arranged in the order from top to bottom and left to right as follows: (a) GAL, (b) BCW, (c) ION, (d) WDB, (e) PID, (f) APP, (g) SAH, (h) HFC, (i) HEP, (j) HCD.

it achieves a standard deviation of merely 0.0047 on WDB and 0.0112 on HEP, reflecting its robust stability. Overall, AIFE-IIFDT exhibits outstanding classification performance on most datasets.

To further examine the convergence behavior and generalization performance of the AIFE-IIFDT model across varying training sample sizes, learning curves were plotted for 12 datasets (as shown in Fig. 6). The horizontal axis represents the number of training samples, while the vertical axes track Accuracy and AUC, illustrating how training and testing performance evolve with increasing sample size. Overall, the model demonstrates good convergence and generalization capabilities across most datasets.

To further examine whether there exist significant differences among the algorithms, the Friedman test was applied using IBM SPSS Statistics 27 to comprehensively compare the errors of the nine algorithms across 12 datasets. The average ranks and test results are presented in Table 8. The AIFE-IIFDT model achieved the minimum average rank of 1.75, followed by XGBoost at 2.83, while GBDT reached a maximum of 6.75. The corresponding p-value for the test statistic was less than 0.001, indicating rejection of the null hypothesis at the 0.05 significance level. This confirms that the Accuracy rates among the eight classifiers exhibit significant differences.

Table 7
Accuracy comparison results for eight classification algorithms.

Datasets		Algorithms							
		AIFE-IIFDT	C4.5	CART	RBFNN	RF	GBDT	XGBoost	IFDT
GAL	Avg	**0.8427**	0.6849	0.7047	0.5786	0.7021	0.6958	0.7823	0.6644
	Std	0.0456	0.0534	**0.0446**	0.5786	0.0785	0.0638	0.0556	0.1079
APP	Avg	**0.9121**	0.8318	0.8212	0.8955	0.9000	0.8061	0.8727	0.8564
	Std	**0.0524**	0.0623	0.0631	0.0734	0.0568	0.0862	0.0591	0.0704
BCW	Avg	**0.8883**	0.7392	0.7083	0.7708	0.7725	0.6833	0.8050	0.7276
	Std	0.0527	0.0611	0.0634	0.0500	0.0425	0.0637	0.0415	**0.0377**
SAH	Avg	0.8513	0.6545	0.6577	**0.8620**	0.6932	0.6459	0.6789	0.6304
	Std	**0.0182**	0.0408	0.0414	0.0410	0.0482	0.0478	0.0354	0.0660
ION	Avg	0.9113	0.9192	0.8915	0.9268	0.9127	0.8878	**0.9324**	0.8454
	Std	0.0457	0.0260	0.0300	0.0380	0.0253	0.0328	**0.0219**	0.0298
HFC	Avg	**0.9067**	0.7722	0.7900	0.8961	0.7744	0.7661	0.8394	0.8028
	Std	**0.0453**	0.0495	0.0507	0.0499	**0.0453**	0.0568	0.0472	0.0672
WDB	Avg	0.9503	0.9301	0.9389	0.9515	0.9550	0.9330	**0.9670**	0.9350
	Std	**0.0047**	0.0229	0.0151	0.0185	0.0158	0.0209	0.0153	0.0408
HEP	Avg	**0.9979**	0.8250	0.7917	0.8521	0.8396	0.7896	0.8542	0.6875
	Std	**0.0112**	0.0933	0.0745	0.0905	0.0659	0.0861	0.0710	0.1008
PID	Avg	**0.9331**	0.7106	0.7182	0.8825	0.6963	0.7335	0.7608	0.7244
	Std	**0.0200**	0.0402	0.0279	0.0382	0.0365	0.0375	0.0260	0.0385
HCD	Avg	**0.9372**	0.7492	0.7503	0.8257	0.8115	0.7426	0.8142	0.7326
	Std	**0.0170**	0.0470	0.0520	0.0443	0.0437	0.0542	0.0503	0.0317
GGC	Avg	**0.8724**	0.8623	0.8558	0.8536	0.7454	0.8425	0.8589	0.8284
	Std	0.0335	0.0269	0.0282	0.0284	0.0505	0.0222	**0.0186**	0.0865
ILP	Avg	0.7359	0.7316	0.6630	**0.9507**	0.7168	0.6729	0.6869	0.7118
	Std	0.0310	0.0304	0.0371	0.0238	0.0460	0.0387	0.0354	**0.0065**

Table 8
Friedman test results for eight classification algorithms.

Algorithms	AIFE-IIFDT	XGBoost	RBFNN	RF	C4.5	CART	IFDT	GBDT
Mean Rank	1.75	2.83	3.00	4.33	5.33	5.58	6.42	6.75
p-value				<0.001				

Table 9
Results of pairwise Friedman tests (Significance Level = 0.05).

Algorithms	AIFE-IIFDT	XGBoost	RBFNN	RF	C4.5	CART	IFDT	GBDT
p-value	-	1.000	1.000	0.274	0.009	0.004	0.000	0.000

Furthermore, the AIFE-IIFDT model was served as the control classification algorithm and conducted pairwise Friedman tests against other classification algorithms. The Bonferroni-corrected p-values are presented in Table 9. The results indicate that the corrected p-values for AIFE-IIFDT versus C4.5, CART, GBDT, and IFDT were less than 0.05, demonstrating that its performance significantly outperforms these algorithms. The corrected p-values for XGBoost, RBFNN, and RF were greater than 0.05, thus failing to achieve statistical significance. This indicates that AIFE-IIFDT performs comparably to these three algorithms. This result further validates the overall competitiveness of the proposed model, particularly demonstrating its strong performance that outperforms ensemble learning methods.

Table 10 compares the average runtime of eight algorithms. AIFE-IIFDT's computational overhead falls between traditional decision trees and ensemble methods on most datasets. For example, it took 58.71 seconds on the GAL, longer than C4.5 and CART but shorter than RF, GBDT, and IFDT. On small-to-medium datasets, its runtime ranges from several seconds to tens of seconds, demonstrating practical application potential. Overall, AIFE-IIFDT model achieves excellent classification performance while maintaining computational overhead within an acceptable range.

4.2.3. Feature selection and interpretability analysis

To evaluate the interpretability of the proposed AIFE-IIFDT model, we first counted the frequency of feature selection across 30 independent replicates. Features with higher selection frequencies were identified as key features for each dataset, and their relevance to the corresponding classification tasks was explored. Table 11 summarizes the critical features across all datasets.

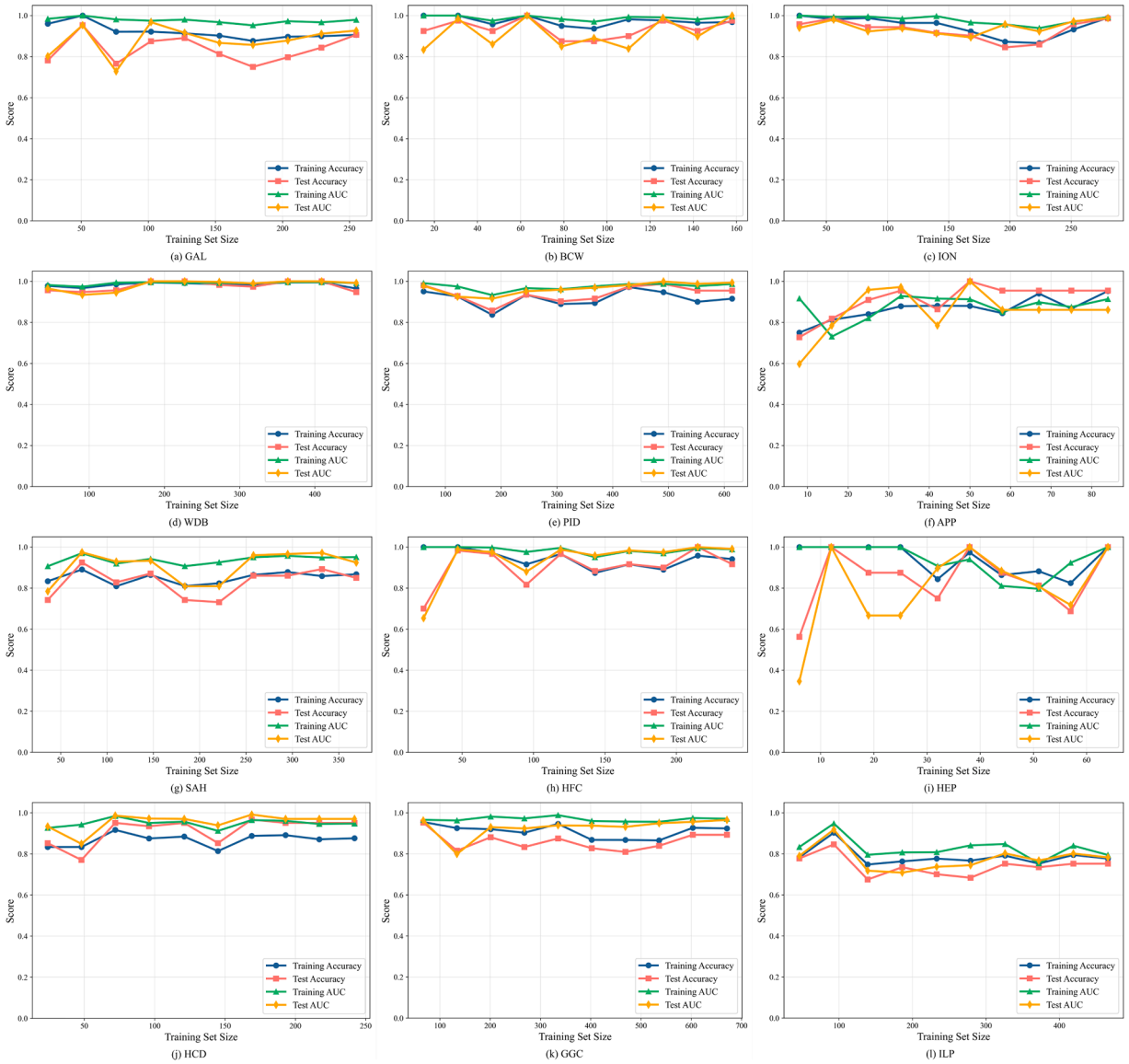


Fig. 6. Accuracy and ROC-AUC learning curve diagram of AIFE-IIFDT. The pictures are arranged in the order from top to bottom and left to right as follows: (a) GAL, (b) BCW, (c) ION, (d) WDB, (e) PID, (f) APP, (g) SAH, (h) HFC, (i) HEP, (j) HCD, (k) GGC, (l) ILP.

Table 10
Comparison of running times for eight classification algorithms (Seconds).

Datasets	Algorithms							
	AIFE-IIFDT	C4.5	CART	RBFNN	RF	GBDT	XGBoost	IFDT
GAL	58.711	0.85	2.04	5.230	350.24	350.39	2.98	1136.63
APP	2.05	0.04	0.10	1.52	24.33	21.39	0.11	444.21
BCW	40.69	0.49	1.65	3.33	220.82	303.47	5.91	2120.54
SAH	16.94	0.37	0.94	6.28	274.17	138.72	1.10	1137.82
ION	24.32	0.92	3.01	6.52	373.08	484.75	7.49	2800.67
HFC	13.92	0.12	0.27	5.27	78.38	53.81	1.93	896.74
WDB	30.73	1.46	4.83	10.32	515.89	853.09	1.62	1983.19
HEP	3.07	0.04	0.05	1.73	14.14	9.62	2.36	664.46
PID	21.49	0.36	1.08	13.15	208.02	164.32	6.62	1439.35
HCD	35.35	0.11	0.23	5.66	73.84	41.36	2.07	1139.92
GGC	35.67	0.53	0.78	24.39	79.48	85.46	144.08	2740.35
ILP	32.12	0.7503	1.32	16.57	285.07	180.68	89.43	480.91

Table 11
Selected features for AIFE-IIFDT across various datasets.

Datasets	Selected Features
GAL	Age, Gender, Comorbidit, Intracellular Water, Extracellular Fluid/Total Body Water, Total Body Fat Ratio, Body Protein Content, Visceral Fat Rating, Bone Mass, Visceral Fat Area, Visceral Muscle Area, Total Cholesterol, High Density Lipoprotein, Aspartat Aminotransferaz, Alkaline Phosphatase, Creatinine, Vitamin D
APP	At1, At2, At3, At4, At5, At6, At7
BCW	Concave_points1, Texture2, Area2, Smoothness2, Concavity2, Smoothness3, Concavity3, Tumor_size, Time
SAH	Sbp, Tobacco, Ldl, Adiposity, Famhist, Typea, Obesity, Alcohol, Age
ION	Attribute1, Attribute3, Attribute5, Attribute7, Attribute16, Attribute22, Attribute29, Attribute34
HFC	Age, Anaemia, Creatinine_phosphokinase, Diabetes, Ejection_fraction, Sex, Smoking, Time, High_blood_pressure, Platelets, Serum_creatinine, Serum_sodium
WDB	Dadius1, Texture1, Concave_points1, Fractal_dimension1, Texture2, Texture3, Perimeter3, Concave_points3
HEP	Steroid, Malaise, Liver Big, Ascites, Varices, Histology
PID	Preg, Plas, Pres, Skin, Insu, Mass, Pedi, Age
HCD	Sex, Trestbps, Chol, Thalach, Oldpeak
GGC	Gende, Age_at_diagnosis, IDH1,IDH2, TP53, ATR, PTEN, EGFR, MUC16, PIK3C,NF1
ILP	Age, Gender, TB, DB, Alkphos, Sgpt, TP, ALB, A/G Ratio

In some datasets, the model tends to select all or most features. For example, in PID, the model relies on all features, which are classic factors for diabetes diagnosis and risk assessment. In SAH, all features were selected, covering the primary modifiable risk factors for cardiovascular disease. In HFC, the model similarly selected all features, which hold clear significance in assessing heart failure prognosis. In ILP, nine routine liver function biochemical indicators including TB, DB, Alkphos, and Sgpt were selected, covering the assessment of hepatic synthetic and metabolic functions.

On other datasets, the model selected only a subset of features. For instance, in the HEP dataset, although only six features were chosen, they all represent key clinical symptoms and pathological indicators for assessing hepatitis severity, directly corresponding to the disease's severity and complications. In HCD, the five features selected by the model are the most commonly used Electrocardiogram and lipid profile indicators for diagnosing heart disease. In GGC, all eleven features relied upon by the model are closely associated with the molecular subtyping and prognosis of glioma, with IDH1/2, TP53, PTEN, and others serving as core biomarkers in the WHO classification of central nervous system tumors.

For BCW and WDB, the features selected by the model are all nuclear morphological measurements, such as concave_points, texture, area, concavity, etc. These features originate from the digital description of breast fine-needle biopsy images and serve as standard parameters for breast cancer diagnosis and prognostic assessment. They correlate with the criteria for assessing malignant nuclear atypia in pathology. In the GAL, the selected features encompass multiple aspects including body composition, lipid metabolism, liver function, and kidney function. These factors have been epidemiologically proven to be directly associated with gallstone formation.

For the anonymized features APP and ION, no in-depth analysis is conducted here as the dataset does not provide the specific physical meaning of these features. In summary, the selected features across all datasets collectively demonstrate the AIFE-IIFDT model's effectiveness in capturing discriminative information and its strong interpretability.

4.2.4. Comparative analysis with existing feature selection methods

The feature selection mechanism of AIFE-IIFDT exhibits fundamental differences from existing methods in both theoretical foundations and selection criteria. Classical feature selection methods are primarily categorized into three types: filtering-based, wrapping-based, and embedding-based. Filter-based methods such as information gain and chi-square tests are independent of learning algorithms. They evaluate feature importance based on statistical relevance, offering computational efficiency but ignoring interactions between features. Wrapper methods such as recursive feature elimination focus on classifier performance, capturing combination effects but incurring significant computational overhead. Embedded methods like Lasso and decision tree intrinsic importance integrate feature selection into model training, achieving feature screening while optimizing objectives and balancing efficiency with effectiveness.

AIFE-IIFDT is an embedded method that employs AIFE as its splitting criterion. Unlike traditional decision trees based on information gain or Gini index, AIFE not only considers the difference between membership and nonmembership but also characterizes sample uncertainty in fuzzy environments through the hesitation component. When splitting nodes, the feature selection process prioritizes those that most significantly reduce overall fuzziness, thereby deeply integrating with intuitionistic fuzzy information processing. Consequently, AIFE-IIFDT model expands the capability of traditional embedded feature selection methods to characterize uncertain information through explicit modeling of intuitionistic fuzzy entropy, offering a unique theoretical advantage when handling complex and fuzzy classification problems.

4.2.5. Computational efficiency and scalability evaluation

To comprehensively evaluate the practical application potential of the proposed AIFE-IIFDT model, this section analyzes it from two dimensions: runtime composition and computational scalability, while exploring its deployment value in scenarios such as medical diagnosis.

(1) Runtime Composition and Scalability Analysis.

Total runtime for AIFE-IIFDT primarily consists of three components: data fuzzification via FCM, parameter optimization, and decision tree construction. The parameter optimization process accounts for the majority of the time, with its computational overhead stemming from grid search traversal of the parameter space. In contrast, both data fuzzification and tree construction can be completed within seconds. This distribution indicates that the computational bottleneck of the model lies in parameter tuning rather than the fundamental modeling steps.

According to the theoretical analysis in Section 3.7, the time complexity of AIFE-IIFDT exhibits a linear relationship with the sample size N, number of features D, and number of classes K. This conclusion is validated by the runtime measurements across 12 datasets (Table 10). The runtime increases with the sample size and feature dimension, maintaining a stable growth trend across different datasets. For the number of classes K, this work focuses on binary classification problems, but the model framework inherently supports extension to multi-class problems while maintaining a linear relationship with K in terms of time complexity.

(2) Potential for Practical Application.

AIFE-IIFDT model demonstrates strengths in fields like medical diagnosis where interpretability is highly demanded. The model outputs classification rules in decision tree format, enabling clinicians to readily grasp the logical relationships between features and diseases. Simultaneously, by explicitly modeling uncertainty, it effectively handles ambiguous information commonly found in medical data, thereby enhancing diagnostic reliability.

For resource-constrained or time-sensitive applications, lightweight strategies can be employed to reduce computational overhead. By increasing the parameter search step size from 0.01 to 0.1, reducing the number of cross-validation folds, or implementing early stopping mechanisms, training time can be compressed to within seconds while maintaining high classification performance.

5. Application of AIFE-IIFDT in thrombus classification

The task of classifying clinical thrombus data is often challenging due to the ambiguity in the relationship between features and disease states, as well as the high dimensionality and incompleteness of the data. Traditional classification models struggle to adequately handle such inherent uncertainty, while existing approaches, such as fuzzy decision trees, remain limited in fully expressing and utilizing hesitation, which is a crucial form of uncertainty information. Therefore, this section introduces the proposed AIFE-IIFDT model into the thrombus classification scenario, aiming to further validate the practicality and robustness of the proposed model through its outstanding performance on real clinical data.

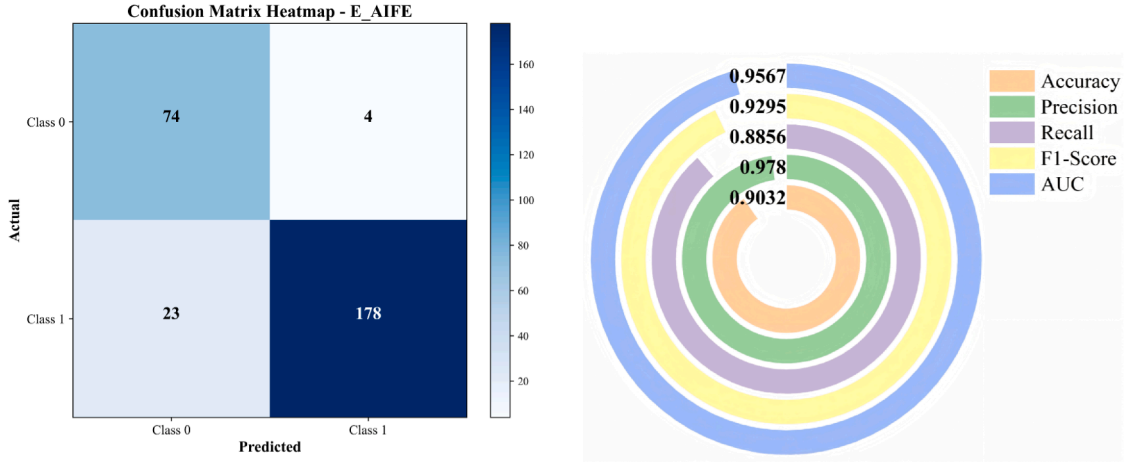
5.1. Dataset description

The study utilized a clinical thrombosis dataset provided by Chongqing Medical University, comprising a total of 1391 samples. Among them, 391 cases were confirmed thrombus patients (194 males, 197 females), and 1000 cases were non-thrombus patients (640 males, 360 females). The dataset comprises 80 features in addition to the target columns (normal and diseased), encompassing red blood cell count, mean corpuscular hemoglobin, D-dimer, creatine kinase isoenzymes, and others, constituting a typical high-dimensional medical dataset.

5.2. Data preprocessing

During the data preprocessing stage, thresholds were first established by reviewing relevant literature and clinical guidelines. Certain raw values were then converted into binary variables: states exceeding the threshold were defined as “abnormal” (1), while those below were defined as “normal” (0). Subsequently, missing values were imputed: for categorical features, the mode was used; for numerical features, the mean was employed. To further enhance the numerical stability of subsequent FCM algorithms and fuzzification processes, robust normalization methods are employed to handle infinite values and NaN. Additionally, minute random noise is introduced to low-variance features to prevent numerical singularities in distance calculations.

During the FCM algorithm process, the following three numerical stability enhancement measures were implemented. These improvements significantly enhanced the algorithm’s robustness and convergence reliability when applied to complex clinical data.



(a) Confusion matrix on the test set (b) Evaluation metrics for thrombus classification

Fig. 7. Classification performance of the AIFE-IIFDT model on the clinical thrombus test dataset. (a) Confusion matrix. (b) Comprehensive evaluation metrics assessing the model's overall performance.

(1) Calculation of the distance: apply a lower bound pruning to the computed Euclidean distance, restricting it to $[\varepsilon, +\infty)$;
(2) Ratio calculation: when calculating the distance ratio, replace outliers with 1.0;
(3) Denominator handling: implement lower bound protection for the denominator in the membership update formula.

5.3. Experimental results and analysis

After completing data preprocessing and enhanced FCM fuzzification, the processed data is fed into the AIFE-IIFDT model for training and evaluation. The training and testing sets were divided in an 8:2 ratio. The optimal hyperparameter combination ($\alpha = 0.28$, $\beta_0 = 0.61$) was determined through grid search and five-fold cross-validation, achieving an average AUC of 0.9814 during cross-validation. As shown in Fig. 7a and b, the confusion matrix and evaluation metrics of the model on the independent test set indicate: among the 279 test samples, the model accurately identified 74 true negatives and 178 true positives, achieving an Accuracy of 0.9032, a Precision of 0.9780, a Recall of 0.8856, an F1-Score of 0.9295, and an AUC of 0.9576.

Among these, false negatives (23 cases) significantly exceeded false positives (4 cases). This phenomenon may stem from two aspects: first, conventional filling methods for missing values fail to fully restore the relationship between characteristics and disease, resulting in the loss of critical discriminative information. Second, the absence of feature selection allowed high-dimensional noise to interfere with the model's precise delineation of thrombus boundaries, leading it to misclassify some atypical thrombus samples as normal.

Experimental results demonstrate that the model exhibits superior comprehensive performance in the thrombus classification task. High accuracy reflects the model's reliability in identifying positive samples, helping to reduce unnecessary testing in clinical practice. A high Recall reflects effective coverage of genuine cases, thereby controlling the risk of missed diagnoses. The F1-Score and AUC validate the model's balanced performance and strong discriminative capability. The empirical research on thrombus clinical data fully demonstrates the practical value of the AIFE-IIFDT model. Not only can it effectively handle uncertainties in high-dimensional medical data, but its high Accuracy and strong robustness also provide reliable solutions for complex medical classification problems, showcasing promising prospects for widespread application.

6. Conclusions and future research directions

Classification problems in the real world often face multiple challenges such as high-dimensional data and noise interference. Traditional models have inherent constraints when handling such uncertain information. To address this issue, this study combines the theory of IFS with decision tree models to propose AIFE and builds upon this foundation to construct the AIFE-IIFDT model. The experimental section, based on multiple independent repetitions, validated the effectiveness of the proposed model by comparing it against five existing intuitionistic fuzzy entropies and eight other classification algorithms across several public datasets. Furthermore, the model has been successfully applied in clinical thrombus classification tasks further demonstrates its practical value in complex medical scenarios.

Despite the aforementioned progress achieved in this study, certain limitations remain. First, the current model is primarily designed for binary classification problems and has not yet been systematically extended to multi-class scenarios. Second, the computational overhead of the model is relatively high during the parameter optimization phase, which may limit its direct application in scenarios involving large-scale data or stringent real-time requirements.

Based on the above analysis, future research may be conducted in the following directions:

(1) Extending to Multi-Class Tasks.

Extending AIFE-IIFDT to multi-class scenarios. Drawing inspiration from approaches to multi-label feature selection and category dependency modeling [49], and establishing an intuitionistic fuzzy decision tree framework for multi-class problems, and exploring explicit modeling methods for inter-category dependencies.

(2) Ensemble Learning and Lightweight Optimization.

Integrating AIFE-IIFDT with ensemble learning frameworks such as RF enhances generalization performance and stability through multi-model fusion. Simultaneously, drawing on the principles of multi-objective meta-heuristic optimization [50], we explore heuristic parameter optimization strategies to reduce the computational overhead of grid search, and design a lightweight version tailored for resource-constrained scenarios.

(3) Extend to More Practical Application Domains.

Apply the proposed model to an expanded range of real-world problems, such as industrial fault diagnosis, text classification, and image recognition, to further validate its generalization capabilities and practicality across different data types and domains.

CRediT authorship contribution statement

Xiaozeng Xu: Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Formal analysis; **Yanni Guo:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation, Conceptualization; **Xing Pu:** Visualization, Data curation; **Weihua Xu:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Data availability

Data will be made available on request.

Declaration of competing interest

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

Acknowledgement

This work was supported by the National Natural Science Foundation of China (NO. 62376229), the Natural Science Foundation of Chongqing (NO. CSTB2023NSCQ-LZX0027).

References

[1] J.W. Weisel, R.I. Litvinov, Visualizing thrombosis to improve thrombus resolution, Res. Pract. Thromb. Haemost. 5 (1) (2021) 38–50.
[2] A. Shukla, S. Giri, Portal vein thrombosis in cirrhosis, J. Clin. Exp. Hepatol. 12 (3) (2022) 965–979.
[3] X. Yang, H. Chen, T. Li, S. Feng, J. Wan, Y. Yao, Adaptive feature selection with weighted fuzzy rough sets for noisy data, Fuzzy Sets Syst. 518 (2025) 109499.
[4] J.R. Quinlan, Induction of decision trees, Mach. Learn. 1 (1) (1986) 81–106.
[5] B. Hssina, A. Merbouha, H. Ezzikouri, M. Erritali, A comparative study of decision tree ID3 and C4.5, Int. J. Adv. Comput. Sci. Appl. 4 (2) (2014) 13–19.
[6] J. Jiang, X. Zhu, G. Han, M. Guizani, L. Shu, A dynamic trust evaluation and update mechanism based on C4.5 decision tree in underwater wireless sensor networks, IEEE Trans. Veh. Technol. 69 (8) (2020) 9031–9040.
[7] L.A. Zadeh, Fuzzy sets, Inf. Contr. 8 (3) (1965) 338–353.
[8] R. Lv, J. Li, Y. Wang, Z. Yang, On entropies of fuzzy sets and nonlinear integrals, Fuzzy Sets Syst. 518 (2025) 109513.
[9] O. Filimonova, A. Ovsyannikov, N. Biryukova, Original conditional hybrid entropy tested on the case of prostate cancer, Fuzzy Sets Syst. 520 (2025) 109560.
[10] C.Z. Janikow, Fuzzy decision trees: issues and methods, IEEE Trans. Syst. Man Cybern. B 28 (1) (1998) 1–14.
[11] M. Umanol, H. Okamoto, I. Hatono, H. Tamura, F. Kawachi, S. Umedzu, J. Kinoshita, Fuzzy decision trees by fuzzy ID3 algorithm and its application to diagnosis systems, in: Proceedings of 1994 IEEE 3rd International Fuzzy Systems Conference, vol. 3, 1994, pp. 2113–2118.
[12] B. Chandra, P.P. Varghese, Fuzzy SLIQ decision tree algorithm, IEEE Trans. Syst. Man Cybern. B 38 (5) (2008) 1294–1301.
[13] B. Chandra, P.P. Varghese, Fuzzifying gini index based decision trees, Expert Syst. Appl. 526 (4) (2009) 8549–8559.
[14] A. Rahimieh, M. Mehriar, S.M. Zamir, M. Nosrati, Fuzzy-decision tree modeling for H2S production management in an industrial-scale anaerobic digestion process, Biochem. Eng. J. 208 (2024) 109380.
[15] Q. Ren, D. Zhang, X. Zhao, L. Yan, J. Rui, et al., A novel hybrid method of lithology identification based on k-means + + algorithm and fuzzy decision tree, J. Pet. Sci. Eng. 208 (2022) 109681.
[16] N.F. Idris, M.A. Ismail, Breast cancer disease classification using fuzzy-ID3 algorithm with FUZZYDBD method: automatic fuzzy database definition, PeerJ Comput. Sci. 7 (2021) e427.
[17] Q. Hu, J. Bai, J. Zhang, Y. Song, J. Mi, FNCEOD: Fuzzy neighborhood combination entropy-based outlier detection, Fuzzy Sets Syst. 526 (2025) 109683.
[18] G. Briglia, F. Immovilli, M. Cocconcelli, M. Lippi, Bearing fault detection and recognition from supply currents with decision trees, IEEE Access 12 (2024) 12760–12770.
[19] K. Pałczyński, C. Magda, T. Talaśka, Fuzzy Gaussian decision tree, J. Comput. Appl. Math. 425 (2023) 115038.
[20] F. Zhu, K. Yu, Y. Lin, C. Wang, J. Wang, M. Chao, Robust LOS/NLOS identification for UWB signals using improved fuzzy decision tree under volatile indoor conditions, IEEE Trans. Instrum. Meas. 72 (2023) 1–11.
[21] L. Jiao, H. Zhang, X. Geng, Q. Pan, Belief rule learning and reasoning for classification based on fuzzy belief decision tree, Int. J. Approx. Reason. 175 (2024) 109300.

[22] X. Zhu, X. Hu, L. Yang, W. Pedrycz, Z. Li, A development of fuzzy-rule-based regression models through using decision trees, IEEE Trans. Fuzzy Syst. 32 (5) (2024) 2976–2986.
[23] K.T. Atanassov, Intuitionistic fuzzy sets, in: Intuitionistic Fuzzy Sets: Theory and Applications, Springer, 1999, pp. 1–137.
[24] P. Burillo, H. Bustince, Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets, Fuzzy Sets Syst. 78 (3) (1996) 305–316.
[25] E. Szmidt, J. Kacprzyk, Entropy for intuitionistic fuzzy sets, Fuzzy Sets Syst. 118 (3) (2001) 467–477.
[26] Q. Jiang, J. Huang, X. Jin, P. Wang, W. Zhou, S. Yao, Medical image fusion using a new entropy measure between intuitionistic fuzzy sets joint Gaussian curvature filter, IEEE Trans. Radiat. Plasma Med. Sci. 7 (5) (2023) 494–508.
[27] S. Kashyap, N. Gandotra, N. Saini, Sonia, Novel entropy measure for intuitionistic fuzzy set for the selection of laptop using COPRAS approach, in: 2024 11th International Conference on Computing for Sustainable Global Development (INDIACom), 2024, pp. 802–807.
[28] L. Zeng, H. Ren, T. Yang, N. Xiong, An intelligent expert combination weighting scheme for group decision making in railway reconstruction, Mathematics 10 (4) (2022) 549.
[29] Y. Lan, Multi-criteria Decision Making Method Based on Intuitionistic Fuzzy Entropy and Its Application (in Chinese), Master's thesis, North Minzu University, 2025.
[30] X. Xu, M. Li, Application of arctangent intuitionistic fuzzy entropy in multi-attribute decision making, Fuzzy Syst. Math. 37 (01) (2023) 49–57.
[31] M. Aggarwal, Bridging the gap between probabilistic and fuzzy entropy, IEEE Trans. Fuzzy Syst. 28 (9) (2020) 2175–2184.
[32] X. Liu, W. Jia, W. Liu, W. Pedrycz, AFSSE: an interpretable classifier with axiomatic fuzzy set and semantic entropy, IEEE Trans. Fuzzy Syst. 28 (11) (2020) 2825–2840.
[33] C. Zhang, Z. Lu, Y. Zhang, J. Dai, Online streaming feature selection using bidirectional complementarity based on fuzzy gini entropy, IEEE Trans. Fuzzy Syst. 33 (5) (2025) 1592–1604.
[34] P. Bujnowski, E. Szmidt, J. Kacprzyk, An approach to intuitionistic fuzzy decision trees, in: 2015 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology (IFSA-EUSFLAT-15), Atlantis Press, 2015, pp. 1253–1260.
[35] Y. Li, Research and Application of Decision Tree Algorithm Based on Intuitionistic Fuzzy Sets (in Chinese), Master's thesis, Beijing Jiaotong University, 2019.
[36] J. Xian, S. Rezvani, D. Yang, A new decision tree based on intuitionistic fuzzy twin support vector machines, IEEE Trans. Intell. Transp. Syst. 25 (12) (2024) 19810–19819.
[37] S.M.S. Askari, M.A. Hussain, IFDTC4.5: Intuitionistic fuzzy logic based decision tree for E-transactional fraud detection, J. Inf. Secur. Appl. 52 (2020) 102469.
[38] T. Djatna, M.K.D. Hardhienata, A.F.N. Masruriyah, An intuitionistic fuzzy diagnosis analytics for stroke disease, J. Big Data 5 (1) (2018) 35.
[39] J.R. Quinlan, Improved use of continuous attributes in C4. 5, J. Artif. Intell. Res. 518 (1996) 77–90.
[40] V. Calabrese, D. Metro, A. Alibrandi, D. Maviglia, V. Cernaro, V. Maressa, E. Longhitano, G. Gembillo, D. Santoro, Review and practical excursus on the comparison between traditional statics methods and classification and regression tree (CART) in real-life data: low protein diet compared to mediterranean diet in patients with chronic kidney disease, Nefrologia 45 (4) (2025) 279–284.
[41] F. Fabris, A. Doherty, D. Palmer, J.P. De Magalhaes, A.A. Freitas, A new approach for interpreting random forest models and its application to the biology of ageing, Bioinformatics 34 (14) (2018) 2449–2456.
[42] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Stat. 29 (2001) 1189–1232.
[43] S.S.M. Ghoneim, M. Baz, A. Alzaed, Y.T. Zewdie, Predicting the insulating paper state of the power transformer based on XGBoost/lightGBM models, Sci. Rep. 15 (1) (2025) 17836.
[44] Q. Que, M. Belkin, Back to the future: radial basis function network revisited, IEEE Trans. Pattern Anal. Mach. Intell. 42 (8) (2019) 1856–1867.
[45] W. Peizhuang, Pattern recognition with fuzzy objective function algorithms (James C. Bezdek), Siam Rev. 25 (3) (1983) 442.
[46] N.R. Pal, J.C. Bezdek, On cluster validity for the fuzzy c-means model, IEEE Trans. Fuzzy Syst. 3 (3) (1995) 370–379.
[47] E. Szmidt, J.F. Baldwin, Intuitionistic fuzzy set functions, mass assignment theory, possibility theory and histograms, in: 2006 IEEE International Conference on Fuzzy Systems, vol. 29, 2006, pp. 35–41.
[48] Y. Ren, X. Zhu, K. Bai, R. Zhang, A new random forest ensemble of intuitionistic fuzzy decision trees, IEEE Trans. Fuzzy Syst. 31 (5) (2023) 1729–1741.
[49] P. Dhal, C. Azad, Zone oriented binary multi-objective charged system search based feature selection approach for multi-label classification, Expert Syst. 42 (2) (2025) e13803.
[50] P. Dhal, B. Pradhan, U. Fiore, S.A.J. Francis, D.S. Roy, A clinical diabetes prediction based support system based on the multi-objective metaheuristic inspired fine tuning deep network, Inf. Fusion 122 (2025) 103188.