An Asymmetric Approach to Three-Way Approximation of Fuzzy Sets

Xuerong Zhao, Duoqian Miao, Yiyu Yao, and Witold Pedrycz, *Life Fellow, IEEE*

***Abstract*—The three-way approximation of fuzzy sets represents membership values using a three-valued set $\{1, \mathrm{m}, 0\}$, where 1 indicates total belongingness, 0 total nonbelongingness, and m an intermediate state. This approach elevates values of membership function above a threshold α to 1, reduces those below β to 0, and assigns the remaining ones to an intermediate value m. A key challenge lies in determining the thresholds α and β and selecting the value of m, as existing models often lack analytical solutions and fail to fully explore the relationship between m and membership structures. This study introduces an asymmetric three-way approximation model for fuzzy sets, removing the constraint $\alpha + \beta = 1$. Analytical formulas are derived for the thresholds α and β by minimizing information loss, and the relationship between m and membership structures is thoroughly examined. An adaptive optimizer is proposed to learn the approximate optimal value of m by minimizing the information loss. The experimental results show that information loss decreases initially before increasing as m grows. Besides, our model achieves the best classification across most datasets.**

***Index Terms*—Fuzzy set, shadowed set, three-way approximation, three-way decision.**

I. Introduction

FUZZY sets extend set theory by allowing partial membership, making them helpful in handling uncertainty and imprecision [1], [2], [3], [4], [5]. This flexibility is advantageous for representing real-world data where boundaries are often unclear. However, while fuzzy sets offer greater expressiveness, their high precision associated with membership values can also hinder interpretability and complicate decision-making processes, as it may be challenging to define appropriate membership functions and decision thresholds.

Received 25 March 2025; accepted 26 April 2025. Date of publication 30 April 2025; date of current version 3 July 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 62006172 and Grant 62376198 and in part by the National Key Research and Development Program of China "Key Special Project on Cyberspace Security Governance" under Grant 2022YFB3104700. Recommended by Associate Editor P. Liu. *(Corresponding author: Duoqian Miao.)*

Xuerong Zhao is with the College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai 201418, China (e-mail: xrzhao@shnu.edu.cn).

Duoqian Miao is with the Department of Computer Science and Technology, Tongji University, Shanghai 201804, China (e-mail: dqmiao@tongji.edu.cn).

Yiyu Yao is with the Department of Computer Science, University of Regina, Regina, SK S4S 0A2, Canada (e-mail: yiyu.yao@uregina.ca).

Witold Pedrycz is with the Department of Measurement and Control Systems, Silesian University of Technology (SUT), 44-100 Gliwice, Poland, also with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2R3, Canada, and also with the Research Center of Performance and Productivity Analysis, Istinye University, Istanbul 34396, Türkiye (e-mail: wpedrycz@ualberta.ca).

Digital Object Identifier 10.1109/TFUZZ.2025.3565700

The three-way approximation of fuzzy sets [6], [7], [8], [9], [10], [11] was introduced to address challenges in interpretability and decision-making associated with high-precision membership functions. It represents a key application of three-way decision theory [12], [13], [14], [15], [16] and can be formally characterized using two distinct approaches: three mutually exclusive sets and a three-valued set.

A. Three Mutually Exclusive Sets

This approach approximates the fuzzy set using three distinct regions: the positive, negative, and boundary regions. The positive region comprises elements with high membership values, indicating they are essential to the fuzzy set. The negative region includes elements with low or zero membership values, signifying they are not part of the set. The boundary region contains elements with intermediate membership values, whose membership status is uncertain or ambiguous. This classification enhances the understanding and interpretation of fuzzy data by clearly tracing the degrees of membership. One way to define these three mutually exclusive sets is by employing a pair of cut sets from fuzzy sets, such as the α-core and β-support [11].

B. Three-Valued Set and Shadowed Set

In this alternative approach, each element of the universe is assigned one of three values, typically denoted as $\{\mathbf{0}, \mathbf{m}, \mathbf{1}\}$. Here, $\mathbf{0}$ signifies total nonmembership (indicating the element is excluded from the set), $\mathbf{1}$ signifies total membership (indicating the element is included in the set), and $\mathbf{m}$ represents an intermediate value that reflects partial or uncertain membership (where the element is neither fully included nor fully excluded). Shadowed sets (SS) are a typical model of three-valued sets and provide a crucial method for constructing three-valued sets from fuzzy sets [6], [7]. This framework simplifies the representation of uncertainty and captures the ambiguity between full membership and nonmembership. The interpretation of these three values can be approached from two distinct perspectives: computational and semantic.

From a *computational perspective*, these values can be assigned specific numerical values to facilitate the model's calculation and application. While $\mathbf{0}$ and $\mathbf{1}$ are straightforward to define, determining the intermediate value $\mathbf{m}$ is more complex and presents a significant challenge. The interpretation of the intermediate value $\mathbf{m}$ in three-way approximations of fuzzy sets varies across different methodologies. One approach represents this uncertainty using an interval. For instance, Pedrycz [6] employed the unit interval [0,1] to capture maximum uncertainty in SS. Building on this, Zhang et al. [17] introduced the interval shadowed sets (ISS), where $\mathbf{m}$ is defined as the interval $[\beta, \alpha]$ ($0 \leq \beta < \alpha \leq 1$). Alternatively, $\mathbf{m}$ can be a specific real number between 0 and 1, offering more flexibility. Cattaneo and Ciucci [18], [19] used 0.5 to represent ambiguous membership grades, leading to 0.5-shadowed sets (0.5-SS). Recognizing the limitations of a fixed value like 0.5, Deng and Yao [20] introduced a mutable constant for $\mathbf{m}$ in their mean-value-based decision-theoretic SS (MVDTSS), which uses the mean of membership grades that are neither 1 nor 0. Gao et al. [21] refined this further by defining $\mathbf{m}$ as the mean of fuzzy entropy, creating mean-entropy-based SS (MESS).

From a *semantic perspective*, these values are assigned meanings carrying certain semantic significance. Yao et al. [10] proposed a generalized three-valued set with values $\mathbf{n}$, $\mathbf{m}$, and $\mathbf{p}$ to represent negative, indeterminate, and positive instances, respectively, providing greater flexibility in modeling uncertainty. Later, Yang and Yao [22], [23] used $\mathbf{w}$, $\mathbf{g}$, and $\mathbf{b}$ to denote three distinct membership grades of SS, representing the white, gray, and black objects of SS.

When developing a three-way approximation model for fuzzy sets, whether using three mutually exclusive sets or a three-valued set, computation of the two thresholds α and β is essential. Elements with membership degrees greater than or equal to α are classified as belonging to the fuzzy set, while those with membership degrees less than or equal to β are excluded. Elements with membership degrees between α and β fall into an intermediate, uncertain category. Thus, determining the appropriate thresholds is critical for constructing an effective three-way approximation model. Typically, three theoretical frameworks are employed to calculate the threshold: optimization, decision, and game theory.

C. Optimization Theory

The optimization-based approach determines thresholds by minimizing or maximizing an objective function, guided by specific principles [6], [7], [8], [10], [17], [20], [21], [24], [25]. Yao et al. [10] outlined three key principles: uncertainty invariance, minimum distance, and least cost. The *uncertainty invariance* principle ensures that the inherent uncertainty of the fuzzy set is preserved during approximation. Pedrycz [6] derived optimal thresholds by minimizing discrepancies in membership values during elevation and reduction operations. Later, Tahayori et al. [24] extended this by accounting for the fuzziness in the shadowed region. Gao et al. [21] introduced a broader optimization approach using fuzzy entropy as the objective. In many cases, α and β are calculated under the assumption that $\alpha + \beta = 1$ to simplify computations. However, this introduces symmetric constraints in the parameters, reducing the model's flexibility. In contrast, the *Minimum Distance* principle aims to minimize the distance between the three-way approximation and the original fuzzy set. Zhou et al. [26] introduced the minimal distance objective function in constructing three-way approximations and concluded that this function is continuous but nonconvex.

D. Decision Theory

The *Least Cost* principle focuses on minimizing the risks or costs associated with decision-making, leading to cost-sensitive methods [9], [27], [28], [29]. This decision-theoretic approach aims to establish thresholds by evaluating various decision criteria. Deng and Yao [9] introduced a decision-theoretic framework for three-way approximations of fuzzy sets, where thresholds are derived based on a loss/cost function. Zhang et al. [27] extended this model by replacing the fixed value 0.5 with a variable value to optimize the thresholds α and β across different data distributions. Ibrahim and William-West [28] further critiqued the rigidity of using 0 and 1 to represent total nonmembership and membership in three-valued sets, proposing a more flexible three-way approximation using a generalized three-valued set $\{\mathbf{n}, \mathbf{m}, \mathbf{p}\}$.

E. Game Theory

The game-theoretic approach leverages strategic interactions to determine threshold values in three-way approximations. Zhang and Yao [30], [31] introduced the game-theoretic SS (GTSS), which computes thresholds by modeling the interplay between elevation and reduction errors as a game, iteratively adjusting threshold values to balance these errors. Zhang et al. [32] extended this approach by incorporating fuzzy entropy into the error characterization, resulting in the fuzzy-entropy-based GTSS (FeGTSS). Gao et al. [33] argued that existing SS rely on a single principle and lack a multiprinciple perspective. As a result, they proposed the UC-GTSS which integrates game analysis between uncertainty and decision cost.

In addition to research on modeling of three-way approximations of fuzzy sets, there is extensive theoretical and applied work in this area [18], [19], [34], [35], [36], [37], [38], [39], [40], [41], [42]. Cattaneo and Ciucci [18], [19] investigated the algebraic structures of 0.5-SS. Zhang et al. [43] extended this by studying three-way approximations of L-fuzzy sets, treating the three values as white, grey, and black memberships. At the same time, the three-way approximation of fuzzy sets has been widely applied in fields such as medicine [44], biology [45], and computer vision [39], [46] to achieve various downstream tasks, including classification [46], [47], [48], clustering [38], [40], [45], [49], [50], [51], [52], image retrieval [39], recommender systems [42], group decision-making [41], and more.

Although research on three-way approximations of fuzzy sets is extensive and encompasses a wide array of methodologies, two fundamental issues continue to pose challenges in the construction of these models.

1) In optimization-based approaches, parameters α and β are often assumed to satisfy $\alpha + \beta = 1$, simplifying computation but imposing symmetry. This assumption limits the model's flexibility in representing real-world complexities, where asymmetric configurations are more appropriate. For example, in medical diagnosis systems, the threshold for accepting a hypothesis (e.g., disease presence) is typically higher than for rejection, resulting in values like $\alpha = 0.85$ and $\beta = 0.1$, violating the symmetry assumption. Such configurations are common in decision risk control and cost-sensitive classification.

2) There is a lack of comprehensive exploration regarding how different values of $\mathbf{m}$ interact with variations in uncertainty and their corresponding effects on three-way approximation results. Understanding this relationship is crucial, as it can significantly enhance the model's robustness and capacity to accommodate varying degrees of uncertainty across different contexts.

To address these challenges, this article introduces an asymmetric three-way approximation model that departs from the conventional requirement that $\alpha + \beta = 1$. This flexibility allows for a more comprehensive representation of uncertainties inherent in decision-making. Additionally, we provide analytical expressions for calculating the optimal thresholds α and β, facilitating easier implementation of the model in practical applications. Moreover, the article delves into the relationship between the intermediate parameter m (which we assume to be a real number within the interval [0,1] and denote as m) and the underlying membership structures, offering insights into how variations in m can influence decision boundaries.

The rest of this article is organized as follows. Section II provides a detailed introduction to the asymmetric three-way approximation model, outlining its theoretical foundations and presenting analytical solutions for the thresholds α, β, and the intermediate value m. Section III conducts different experiments to demonstrate the model's effectiveness and to address the limitations associated with using fixed values, such as 0.5, for the parameter m. Finally, Section IV concludes this article.

II. Framework for Asymmetric Three-Way Approximation

In this section, we present the asymmetric three-way approximation model, which relaxes the traditional symmetric constraints on threshold values, allowing for flexibility in defining the boundaries between different regions. This approach enables a more subtle treatment of uncertainty by decoupling the thresholds α and β, making them independently adjustable. We also provide analytical solutions for determining the optimal threshold values, offering precise calculations for α and β in this context. Additionally, we explore the role of the intermediate parameter m and its relationship with the underlying membership structures, shedding light on how it influences the classification process within the model.

A. Asymmetric Three-Way Approximation Approach

Let A be a fuzzy set with membership function $\mu_A : U \to [0,1]$, where $U = \{x_1, x_2, \ldots, x_n\}$ represents the universal set of objects. By introducing threshold $(\alpha, \beta) \in [0,1]^2$ with $\beta < \alpha$, we approximated A with a three-valued set

$$T^m_{\alpha,\beta}(A) = \begin{cases} 1, & \mu_A(x_i) \ge \alpha \\ m, & \beta < \mu_A(x_i) < \alpha \\ 0, & \mu_A(x_i) \le \beta \end{cases}$$

where $m \in (0,1)$ is an intermediate value. The approximation method focuses on the membership degrees of elements in a fuzzy set rather than the elements themselves. These membership values, ordered in ascending or descending order, can be represented as discrete points in a Cartesian coordinate system. Each element x_i is mapped to a pair $(\mu_A(x_i), \mu_A(x_i))$, where both coordinates reflect its membership degree in fuzzy set A. This geometric representation uses the line function $f(x) = x$ or $f(x) = -x$ (with $x \in [0,1]$) as a reference to model membership degree distribution, enabling analysis of approximations within the unit square domain while preserving the structural information of membership degrees.

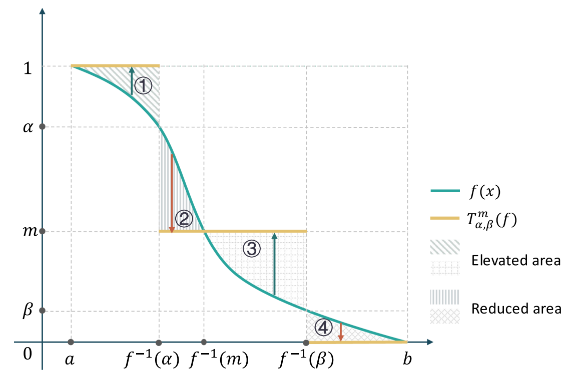

Fig. 1. Three-way approximation of f.

To generalize, let $f : [a,b] \to [0,1]$ be a continuously differentiable, strictly decreasing function. Given that $\alpha, \beta, m \in [0,1]$ satisfy $0 \le \beta < m < \alpha \le 1$. The asymmetric three-way approximation (A3WA for short) of f is defined as

$$T^m_{\alpha,\beta}(f) = \begin{cases} 1, & f(x) \ge \alpha \\ m, & \beta < f(x) < \alpha \\ 0, & f(x) \le \beta. \end{cases} \tag{1}$$

Here, "asymmetric" indicates that $\alpha + \beta \ne 1$. The three-way approximation represents a qualification of f: values greater than or equal to α are mapped to 1, those less than or equal to β are mapped to 0, and values in between are mapped to the intermediate level m, reflecting uncertainty. Fig. 1 illustrates this process, where upward arrows indicate regions where values are raised to 1 and m, and downward arrows show where values are lowered to 0 and m. Regions marked ① and ④ are certain, while ② and ③ represent uncertain regions with values approximated as m.

The key challenge in A3WA is determining the optimal thresholds α, β, and m. Various methods have been proposed over the past two decades, each with its strengths and limitations depending on the application scenario [10]. However, a consensus on the best approach remains lacking. In the following, we establish a relationship among the thresholds by minimizing information loss, defined as the reduction in membership information.

B. Threshold Optimization for Information Loss Minimization

When a value increases from a to b, the information loss is $b - a$ units; when a value decreases from c to d, the information loss is $c - d$ units. Therefore, the total information loss from the three-way approximation of f can be computed as follows:

Definition 1 (Information loss for A3WA): Let $f : [a, b] \to [0, 1]$ be a continuously differentiable and strictly decreasing function. Given that $\alpha, \beta, m \in [0, 1]$ satisfy $0 \le \beta < m < \alpha \le 1$. The information loss for the three-way approximation of f is defined as

$$L(\alpha,\beta,m) = \lambda_1 \int_a^{f^{-1}(\alpha)} (1 - f(x))\mathrm{d}x + \lambda_2 \int_{f^{-1}(\alpha)}^{f^{-1}(m)} (f(x) - m)\mathrm{d}x + \mu_1 \int_{f^{-1}(\beta)}^{b} f(x)\mathrm{d}x + \mu_2 \int_{f^{-1}(m)}^{f^{-1}(\beta)} (m - f(x))\mathrm{d}x \tag{2}$$

where $\lambda_1, \lambda_2, \mu_1, \mu_2$ are nonnegative factors that balance the membership loss in certain and uncertain regions.

The factors λ_1, λ_2, μ_1, and μ_2 in (2) quantify distinct information losses in the three-way approximation. Specifically, λ_1 penalizes assigning elements with membership degrees not less than α to the positive region, accounting for residual uncertainty; λ_2 addresses the loss from assigning elements in (m, α) to the uncertain region, reflecting deviation from the neutral threshold m. Likewise, μ_1 penalizes assigning elements with membership degrees not greater than β to the negative region, capturing uncertainty in rejection, and μ_2 corresponds to the loss from assigning elements in (β, m) to the uncertain region, representing hesitation in rejecting low-support elements.

For convenience, we denote

$$L(\alpha,m) = \lambda_1 \int_a^{f^{-1}(\alpha)} (1 - f(x))\mathrm{d}x + \lambda_2 \int_{f^{-1}(\alpha)}^{f^{-1}(m)} (f(x) - m)\mathrm{d}x \tag{3}$$

$$L(\beta,m) = \mu_1 \int_{f^{-1}(\beta)}^{b} f(x)\mathrm{d}x + \mu_2 \int_{f^{-1}(m)}^{f^{-1}(\beta)} (m - f(x))\mathrm{d}x. \tag{4}$$

Thus, the total information loss can be rewritten as

$$L(\alpha,\beta,m) = L(\alpha,m) + L(\beta,m). \tag{5}$$

Since $L(\alpha,m) \ge 0$ and $L(\beta,m) \ge 0$, minimizing $L(\alpha,\beta,m)$ for a fixed m is equivalent to independently minimizing $L(\alpha,m)$ and $L(\beta,m)$. This is formalized in the following theorem, which provides explicit formulas for the optimal values α^{opt} and β^{opt}, assuming f is continuously differentiable and strictly decreasing.

Theorem 1: Let $f : [a, b] \to [0, 1]$ be a continuously differentiable and strictly decreasing function, and $m \in [0, 1]$ be a fixed intermediate value. The total information loss function $L(\alpha,\beta,m)$ is separable as

$$L(\alpha,\beta,m) = L(\alpha,m) + L(\beta,m)$$

with $L(\alpha,m)$ and $L(\beta,m)$ depending only on α and β, respectively. The optimal values α^{opt} and β^{opt} are found by solving the independent subproblems

$$\alpha^{\text{opt}} = \arg\min_{\alpha} L(\alpha,m) \quad \beta^{\text{opt}} = \arg\min_{\beta} L(\beta,m).$$

These optimal values are explicitly given by

$$\alpha^{\text{opt}} = \frac{\lambda_1 + \lambda_2 m}{\lambda_1 + \lambda_2} \quad \beta^{\text{opt}} = \frac{\mu_2 m}{\mu_1 + \mu_2}. \tag{6}$$

Moreover, $(\alpha^{\text{opt}}, \beta^{\text{opt}})$ minimizes the total information loss function, satisfying

$$(\alpha^{\text{opt}}, \beta^{\text{opt}}) = \arg\min_{(\alpha,\beta)} L(\alpha,\beta,m). \tag{7}$$

Proof: We aim to find the unique optimal values α^{opt} and β^{opt} that minimize $L(\alpha,m)$ and $L(\beta,m)$, respectively, and show that these solutions achieve the global minimum of $L(\alpha,\beta,m)$.

Step 1. Optimization for α^{opt}: Taking the derivative of $L(\alpha,m)$ [shown in (3)] for α, we have

$$\frac{\partial}{\partial\alpha} L(\alpha,m) = \frac{\lambda_1(1-\alpha) - \lambda_2(\alpha - m)}{f'(f^{-1}(\alpha))}.$$

Setting $\frac{\partial}{\partial\alpha} L(\alpha,m) = 0$, we solve $\lambda_1(1-\alpha) = \lambda_2(\alpha - m)$. This simplifies to

$$\alpha = \frac{\lambda_1 + \lambda_2 m}{\lambda_1 + \lambda_2}.$$

To ensure the solution is the global minimum, note that $f'(f^{-1}(\alpha)) < 0$ and

1) $\frac{\partial}{\partial\alpha} L(\alpha,m) < 0$ for $\alpha < \frac{\lambda_1+\lambda_2 m}{\lambda_1+\lambda_2}$
2) $\frac{\partial}{\partial\alpha} L(\alpha,m) > 0$ for $\alpha > \frac{\lambda_1+\lambda_2 m}{\lambda_1+\lambda_2}$.

Thus, $L(\alpha,m)$ has a unique critical point at $\alpha^{\text{opt}} = \frac{\lambda_1+\lambda_2 m}{\lambda_1+\lambda_2}$, which is a global minimum.

Step 2. Optimization for β^{opt}: Taking the derivative of $L(\beta,m)$ [shown in (4)] with respect to β, we have

$$\frac{\partial}{\partial\beta} L(\beta,m) = \frac{-\mu_1\beta + \mu_2(m - \beta)}{f'(f^{-1}(\beta))}.$$

Setting $\frac{\partial}{\partial\beta} L(\beta,m) = 0$, we solve $-\mu_1\beta + \mu_2(m-\beta) = 0$. This simplifies to

$$\beta = \frac{\mu_2 m}{\mu_1 + \mu_2}.$$

As with α, note that $f'(f^{-1}(\beta)) < 0$ and

1) $\frac{\partial}{\partial\beta} L(\beta,m) < 0$ for $\beta < \frac{\mu_2 m}{\mu_1+\mu_2}$
2) $\frac{\partial}{\partial\beta} L(\beta,m) > 0$ for $\beta > \frac{\mu_2 m}{\mu_1+\mu_2}$.

Thus, $L(\beta,m)$ has a unique critical point at $\beta^{\text{opt}} = \frac{\mu_2 m}{\mu_1+\mu_2}$, which is a global minimum.

Step 3 Uniqueness and globality of $L(\alpha,\beta,m)$: The total loss $L(\alpha,\beta,m)$ is given by

$$L(\alpha,\beta,m) = L(\alpha,m) + L(\beta,m).$$

Since $L(\alpha,m)$ and $L(\beta,m)$ are minimized independently at their respective unique global minima, substituting α^{opt} and β^{opt} ensures that $L(\alpha,\beta,m)$ is globally minimized. The separability of the terms guarantees that solving the subproblems independently leads to the global minimum for the total loss. ■

This proof ensures that the solutions for α and β are unique and globally optimal. Equation (6) demonstrates the relationship between the optimal values of α and β and the intermediate value m. This relationship is independent of the specific form of the function f, suggesting that the results possess a general characteristic.

Note 1: Equation (6) suggests that imposing the condition $\alpha + \beta = 1$ is not necessary, although this condition holds when the parameters λ_1, λ_2, μ_1, μ_2, and the intermediate value m satisfy the equation

$$m = \frac{\lambda_2\mu_1 + \lambda_2\mu_2}{\lambda_1\mu_2 + \lambda_2\mu_1 + 2\lambda_2\mu_2}$$

such as when $\lambda_1 = \lambda_2$, $\mu_1 = \mu_2 = 1$, and $m = 0.5$. Thus, insisting on $\alpha + \beta = 1$ is unnecessary, though it can simplify solving the optimization problem in (7). Furthermore, when $\lambda_1 = \mu_1 = \lambda$, $\lambda_2 = \mu_2 = 1$, and $m = 0.5$, the result is $(\alpha, \beta) = (\frac{0.5+\lambda}{1+\lambda}, \frac{0.5}{1+\lambda})$, which matches the result presented by Yue et al. [47].

The following outlines several properties of the optimal values α^{opt} and β^{opt}.

Proposition 1: The optimal thresholds α^{opt} and β^{opt} exhibit the following monotonicity property:

1) α^{opt} monotonically increases with λ_1 and monotonically decreases with λ_2;
2) β^{opt} monotonically decreases with μ_1 and monotonically increases with μ_2;
3) α^{opt} and β^{opt} monotonically increase with m.

Proof:

1) Taking the derivative of α^{opt} with respect to λ_1 gives $\frac{\partial \alpha^{\text{opt}}}{\partial \lambda_1} = \frac{\lambda_2(1-m)}{(\lambda_1+\lambda_2)^2}$. Since $m \in [0, 1]$ and $\lambda_2 \geq 0$, we have $\frac{\partial \alpha^{\text{opt}}}{\partial \lambda_1} \geq 0$, implying that α^{opt} increases monotonically with λ_1. Similarly, taking the derivative of α^{opt} with respect to λ_2 yields $\frac{\partial \alpha^{\text{opt}}}{\partial \lambda_2} = \frac{\lambda_1(m-1)}{(\lambda_1+\lambda_2)^2}$. Using the same reasoning, we have $\frac{\partial \alpha^{\text{opt}}}{\partial \lambda_2} \leq 0$, indicating that α^{opt} decreases as λ_2 increases.
2) This is similarly proved.
3) This result is straightforward and self-evident.

■

As a result of Proposition 1, we observe that the shadow area increases monotonically with λ_1 and μ_1, while it decreases monotonically with λ_2 and μ_2. Additionally, the value of m affects the positions of α and β—increasing m leads to higher values of both α and β. The above discussion assumes a fixed value of m. Next, we outline the method for determining the optimal value of m for a given pair of α and β.

Theorem 2: Let $f : [a, b] \to [0, 1]$ be a continuously differentiable and strictly decreasing function. For a fixed pair (α, β) satisfying $0 \leq \beta \leq \alpha \leq 1$, the optimal value of m in the optimization problem

$$m^{\text{opt}} = \arg\min_{m} L(\alpha, \beta, m)$$

is given by

$$m^{\text{opt}} = f\left(\frac{\lambda_2 f^{-1}(\alpha) + \mu_2 f^{-1}(\beta)}{\lambda_2 + \mu_1}\right). \tag{8}$$

Proof: We begin by differentiating the loss function $L(\alpha, \beta, m)$ concerning m, as described in (2). This yields

$$\frac{\partial}{\partial m} L(\alpha, \beta, m)$$
$$= \lambda_2(f^{-1}(\alpha) - f^{-1}(m)) + \mu_2(f^{-1}(\beta) - f^{-1}(m)).$$

Setting $\frac{\partial}{\partial m} L(\alpha, \beta, m) = 0$, we solve

$$f^{-1}(m) = \frac{\lambda_2 f^{-1}(\alpha) + \mu_2 f^{-1}(\beta)}{\lambda_2 + \mu_2}.$$

From this, we can express m as

$$m = f\left(\frac{\lambda_2 f^{-1}(\alpha) + \mu_2 f^{-1}(\beta)}{\lambda_2 + \mu_2}\right).$$

Next, we examine the second derivative of $L(\alpha, \beta, m)$ with m. Differentiating $\frac{\partial}{\partial m} L(\alpha, \beta, m)$ with m, we get

$$\frac{\partial^2}{\partial m^2} L(\alpha, \beta, m) = -\frac{\lambda_2 + \mu_2}{f'(f^{-1}(m))}.$$

Since f is strictly decreasing, $f'(x) < 0$ for all $x \in [a, b]$. Therefore, the second derivative is always positive, namely, $\frac{\partial^2}{\partial m^2} L(\alpha, \beta, m) > 0$. This implies that $L(\alpha, \beta, m)$ is convex with m. Thus, the critical point is the global maximum. Therefore, the optimal value of m is uniquely determined by

$$m^{\text{opt}} = f\left(\frac{\lambda_2 f^{-1}(\alpha) + \mu_2 f^{-1}(\beta)}{\lambda_2 + \mu_2}\right).$$

■

The theorem shows that the optimal intermediate value m^{opt} is the function value of a weighted average of the inverse images of α and β under f^{-1}, with weights λ_2 and μ_2. This provides a solid mathematical foundation for determining m based on α and β. In practice, m is selected according to task-specific criteria. Once m is defined, α and β are efficiently computed using (6), ensuring minimal information loss.

Note 2: If f is a continuously differentiable and strictly increasing function, the principal results remain consistent with those in Theorems 1 and 2. Therefore, further discussion on monotonically increasing continuous functions is omitted to avoid redundancy.

Note 3: The information loss for a discrete membership function is computed as

$$\begin{aligned} L(\alpha, \beta, m) = \lambda_1 &\sum_{\{x|\alpha \leq f(x) \leq 1\}} (1 - f(x)) \\ &+ \lambda_2 \sum_{\{x|m \leq f(x) < \alpha\}} (f(x) - m) \\ &+ \mu_1 \sum_{\{x|0 \leq f(x) \leq \beta\}} f(x) \\ &+ \mu_2 \sum_{\{x|\beta < f(x) < m\}} (m - f(x)) \end{aligned}$$

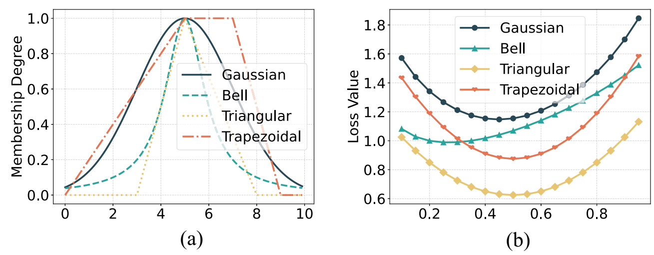

Fig. 2. Illustration of membership functions and information loss with varying parameters. (a) Membership Functions. (b) Information Loss vs. m Variations.

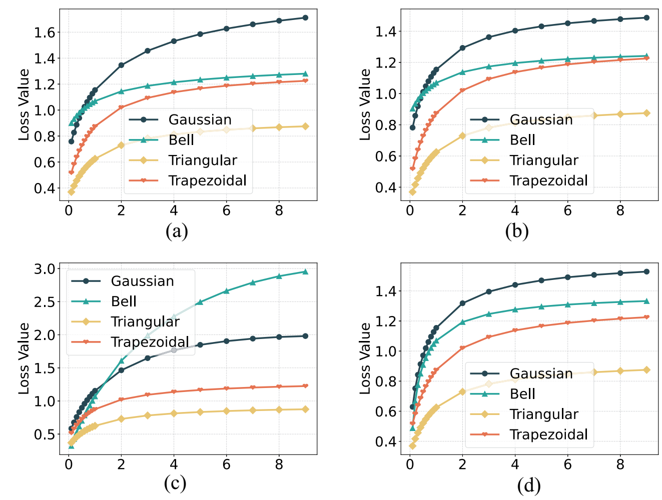

Fig. 3. Illustration of membership functions and information loss with varying parameters. (a) Information Loss vs. λ_1 Variations. (b) Information Loss vs. λ_2 Variations. (c) Information Loss vs. μ_1 Variations. (d) Information Loss vs. μ_2 Variations.

where α and β are computed by (6).

III. Experimental Studies

This section presents experiments to evaluate the proposed A3WA model. We start with a sensitivity analysis to examine how different parameters affect information loss, using both continuous membership functions and UCI datasets. The second part compares classification performance, evaluating the A3WA model against others with various classifiers. We conduct comparative experiments and statistical tests and analyze the impact of parameters on accuracy, providing insights into the model's effectiveness.

A. Sensitivity Analysis of Parameters Affecting Information Loss

In this section, we perform two parametric sensitivity analyses, one based on continuous membership functions and the other on discrete data. These analyses explore the relationship between the intermediate value and membership structures and the impact of each parameter's variation on information loss.

1) Continuous Membership Function Analysis: This section conducts a parametric sensitivity analysis of commonly used membership functions, including Gaussian, bell, triangular, and trapezoidal functions. The membership functions are defined as follows, with Fig. 2(a) showing visual representations of each function.

1) *Gaussian function:*

$$f(x) = e^{-\frac{(x-c)^2}{2\theta^2}}, \quad (\theta, c) = (2, 5).$$

2) *Bell function:*

$$f(x) = \frac{1}{1 + \left|\frac{x-c}{a}\right|^{2b}}, \quad (a, b, c) = (1, 1, 5).$$

3) Triangular function:

$$f(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x \le b, \\ \frac{c-x}{c-b}, & b < x \le c \\ 0, & x > c. \end{cases} \quad (a, b, c) = (3, 5, 8)$$

4) *Trapezoidal function:*

$$f(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1, & b < x \le c, \\ \frac{d-x}{d-c}, & c < x \le d \\ 0, & x > d. \end{cases} \quad (a, b, c, d) = (0, 5, 7, 9)$$

We conducted five experiments to analyze the impact of the intermediate value m and the factors $\lambda_1, \lambda_2, \mu_1, \mu_2$ on information loss. In each experiment, we kept the other four parameters fixed and varied the selected parameter according to predefined rules: $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \{0.1, 0.2, \ldots, 1.0, 2.0, \ldots, 9.0\}$, while m took values from $\{0.1, 0.15, \ldots, 0.95\}$. The fixed values for $\lambda_1, \lambda_2, \mu_1, \mu_2$ and m are $1.0, 1.0, 1.0, 1.0$, and 0.5.

Fig. 2(b) shows how information loss changes with m for each function. Initially, information loss decreases as m increases, reaches a minimum, and then increases. Table I lists the thresholds and corresponding losses. As m increases, both α and β increase. Notably, $m = 0.5$ is not always optimal. The optimal value for the Gaussian function is 0.45; for the bell function, it is 0.25; for the triangular and trapezoidal functions, it is 0.5. Additionally, $m = 0.5$ is the only value for which $\alpha + \beta = 1$, corresponding to the S3WA model in [47] and [48].

Fig. 3(a)–(d) illustrates that the information loss increases with $\lambda_1, \lambda_2, \mu_1, \mu_2$ across all membership functions. Detailed results are provided in Tables II–V. The information loss is constant for all tables involving triangular and trapezoidal functions. This behavior arises from the parameter settings in the second and third experiments.

1) *Second experiment:*
$\lambda_1 \in \{0.1, 0.2, \ldots, 1.0, 2.0, \ldots, 9.0\}$, $\lambda_2 = \mu_1 = \mu_2 = 1.0$, $m = 0.5$.

2) *Third experiment:*
$\lambda_2 \in \{0.1, 0.2, \ldots, 1.0, 2.0, \ldots, 9.0\}$, $\lambda_1 = \mu_1 = \mu_2 = 1.0$, $m = 0.5$.

TABLE I
INFORMATION LOSS VERSUS m VARIATIONS

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
m	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
α	0.550	0.575	0.600	0.625	0.650	0.675	0.700	0.725	0.750	0.775	0.800	0.825	0.850	0.875	0.900	0.925	0.950	0.975
β	0.050	0.075	0.100	0.125	0.150	0.175	0.200	0.225	0.250	0.275	0.300	0.325	0.350	0.375	0.400	0.425	0.450	0.475
Gaussian func.	1.5718	1.4410	1.3416	1.2667	1.2120	1.1750	1.1538	**1.1472**	1.1543	1.1745	1.2078	1.2539	1.3132	1.3861	1.4737	1.5774	1.6999	1.8461
Bell func.	1.0823	1.0278	0.9999	**0.9890**	0.9900	1.0001	1.0174	1.0408	1.0693	1.1025	1.1399	1.1813	1.2267	1.2761	1.3297	1.3879	1.4516	1.5223
Triangular func.	1.0250	0.9313	0.8500	0.7813	0.7250	0.6813	0.6500	0.6313	**0.6250**	0.6312	0.6500	0.6812	0.7250	0.7812	0.8500	0.9313	1.0250	1.1313
Trapezoidal func.	1.4350	1.3038	1.1900	1.0937	1.0150	0.9537	0.9100	0.8837	**0.8750**	0.8837	0.9100	0.9537	1.0150	1.0938	1.1900	1.3038	1.4350	1.5838

Bold values indicate the minimum information loss among the varying values of m.

TABLE II
INFORMATION LOSS VERSUS λ_1 VARIATIONS

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
λ_1	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00
α	0.545	0.583	0.615	0.643	0.667	0.688	0.706	0.722	0.737	0.750	0.833	0.875	0.900	0.917	0.929	0.938	0.944	0.950
β	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250
Gaussian func.	**0.7581**	0.8272	0.8868	0.9388	0.9847	1.0257	1.0625	1.0959	1.1264	1.1543	1.3460	1.4566	1.5308	1.5850	1.6269	1.6604	1.6881	1.7115
Bell func.	**0.8994**	0.9309	0.9573	0.9798	0.9994	1.0167	1.0320	1.0457	1.0581	1.0693	1.1445	1.1865	1.2141	1.2342	1.2495	1.2618	1.2718	1.2803
Triangular func.	**0.3693**	0.4167	0.4567	0.4911	0.5208	0.5469	0.5699	0.5903	0.6086	0.6250	0.7292	0.7813	0.8125	0.8333	0.8482	0.8594	0.8681	0.8750
Trapezoidal func.	**0.5170**	0.5833	0.6394	0.6875	0.7292	0.7656	0.7978	0.8264	0.8520	0.8750	1.0208	1.0938	1.1375	1.1667	1.1875	1.2031	1.2153	1.2250

Bold values indicate the minimum information loss among the varying values of λ_1.

TABLE III
INFORMATION LOSS VERSUS λ_2 VARIATIONS

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$\lambda 2$	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00
α	0.955	0.917	0.885	0.857	0.833	0.812	0.794	0.778	0.763	0.750	0.667	0.625	0.600	0.583	0.571	0.562	0.556	0.550
β	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250	0.250
Gaussian func.	**0.7818**	0.8580	0.9187	0.9688	1.0111	1.0475	1.0793	1.1072	1.1320	1.1543	1.2932	1.3621	1.4035	1.4312	1.4510	1.4659	1.4774	1.4867
Bell func.	**0.9035**	0.9355	0.9617	0.9837	1.0027	1.0193	1.0339	1.0470	1.0587	1.0693	1.1380	1.1738	1.1958	1.2108	1.2217	1.2299	1.2364	1.2416
Triangular func.	**0.3693**	0.4167	0.4567	0.4911	0.5208	0.5469	0.5699	0.5903	0.6086	0.6250	0.7292	0.7813	0.8125	0.8333	0.8482	0.8594	0.8681	0.8750
Trapezoidal func.	**0.5170**	0.5833	0.6394	0.6875	0.7292	0.7656	0.7978	0.8264	0.8520	0.8750	1.0208	1.0938	1.1375	1.1667	1.1875	1.2031	1.2153	1.2250

Bold values indicate the minimum information loss among the varying values of λ_2.

TABLE IV
INFORMATION LOSS VERSUS μ_1 VARIATIONS

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
μ_1	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00
α	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750
β	0.455	0.417	0.385	0.357	0.333	0.312	0.294	0.278	0.263	0.250	0.167	0.125	0.100	0.083	0.071	0.062	0.056	0.050
Gaussian func.	**0.5838**	0.6764	0.7584	0.8318	0.8980	0.9583	1.0134	1.0640	1.1108	1.1543	1.4640	1.6473	1.7667	1.8480	1.9039	1.9419	1.9664	1.9807
Bell func.	**0.3214**	0.4259	0.5234	0.6148	0.7009	0.7823	0.8594	0.9327	1.0026	1.0693	1.6073	1.9900	2.2760	2.4944	2.6622	2.7900	2.8854	2.9537
Triangular func.	**0.3693**	0.4167	0.4567	0.4911	0.5208	0.5469	0.5699	0.5903	0.6086	0.6250	0.7292	0.7813	0.8125	0.8333	0.8482	0.8594	0.8681	0.8750
Trapezoidal func.	**0.5170**	0.5833	0.6394	0.6875	0.7292	0.7656	0.7978	0.8264	0.8520	0.8750	1.0208	1.0938	1.1375	1.1667	1.1875	1.2031	1.2153	1.2250

Bold values indicate the minimum information loss among the varying values of μ_1.

TABLE V
INFORMATION LOSS VERSUS μ_2 VARIATIONS

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
μ_2	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00
α	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750	0.750
β	0.045	0.083	0.115	0.143	0.167	0.187	0.206	0.222	0.237	0.250	0.333	0.375	0.400	0.417	0.429	0.438	0.444	0.450
Gaussian func.	**0.6289**	0.7520	0.8423	0.9132	0.9710	1.0193	1.0604	1.0959	1.1269	1.1543	1.3180	1.3951	1.4402	1.4698	1.4908	1.5065	1.5186	1.5282
Bell func.	**0.4875**	0.6656	0.7742	0.8504	0.9079	0.9532	0.9901	1.0209	1.0469	1.0693	1.1934	1.2467	1.2766	1.2958	1.3092	1.3191	1.3267	1.3327
Triangular func.	**0.3693**	0.4167	0.4567	0.4911	0.5208	0.5469	0.5699	0.5903	0.6086	0.6250	0.7292	0.7812	0.8125	0.8333	0.8482	0.8594	0.8681	0.8750
Trapezoidal func.	**0.5170**	0.5833	0.6394	0.6875	0.7292	0.7656	0.7978	0.8264	0.8520	0.8750	1.0208	1.0938	1.1375	1.1667	1.1875	1.2031	1.2153	1.2250

Bold values indicate the minimum information loss among the varying values of μ_2.

Under these settings, the optimal values $\alpha^{\text{opt}}(\lambda_1)$ and $\alpha^{\text{opt}}(\lambda_2)$ are given by

$$\alpha^{\text{opt}}(\lambda_1) - 0.5 = \frac{\lambda_1 + 1 \cdot 0.5}{\lambda_1 + 1} - 0.5 = \frac{0.5\lambda_1}{1+\lambda_1}$$

$$1 - \alpha^{\text{opt}}(\lambda_2) = 1 - \frac{1 - \lambda_2 \cdot 0.5}{1 + \lambda_2} = \frac{0.5\lambda_2}{1+\lambda_2}.$$

Since λ_1 and λ_2 vary over the same range of values, it follows that $\alpha^{\text{opt}}(\lambda_1) - 0.5 = 1 - \alpha^{\text{opt}}(\lambda_2)$. This symmetry ensures that the area of the green triangle in Fig. 4(a) equals that in Fig. 4(b), and similarly for the orange triangles. This relationship holds only for linear functions. Since $\lambda_1 \int_a^{f^{-1}(\alpha^{\text{opt}}(\lambda_1))}(1 - f(x))\mathrm{d}x + \int_{f^{-1}(\alpha^{\text{opt}}(\lambda_1))}^{f^{-1}(0.5)}(f(x) - 0.5)\mathrm{d}x = \int_a^{f^{-1}(\alpha^{\text{opt}}(\lambda_2))}(1 - f(x))\mathrm{d}x + \lambda_2 \int_{f^{-1}(\alpha^{\text{opt}}(\lambda_1))}^{f^{-1}(0.5)}(f(x) - 0.5)\mathrm{d}x$. We have $L(\alpha^{\text{opt}}(\lambda_1), 0.5) = L(\alpha^{\text{opt}}(\lambda_2), 0.5)$. Thus, the total information loss is identical in the second and third experiments. Since $\mu_1 = \mu_2 = 1$, this equivalence also holds for $L(\beta, 0.5)$. Similar conclusions apply to the fourth and fifth experiments.

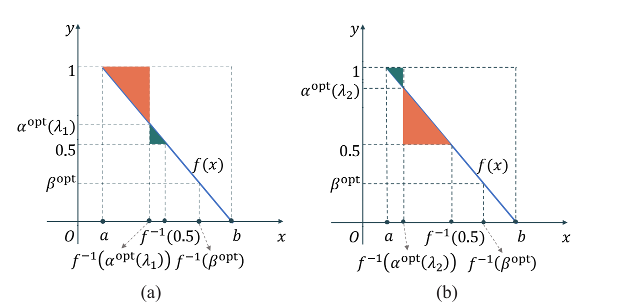

Fig. 4. Illustration of information loss computation. (a) $\lambda_2 = \mu_1 = \mu_2 = 1.0$, $m = 0.5$. (b) $\lambda_1 = \mu_1 = \mu_2 = 1.0$, $m = 0.5$

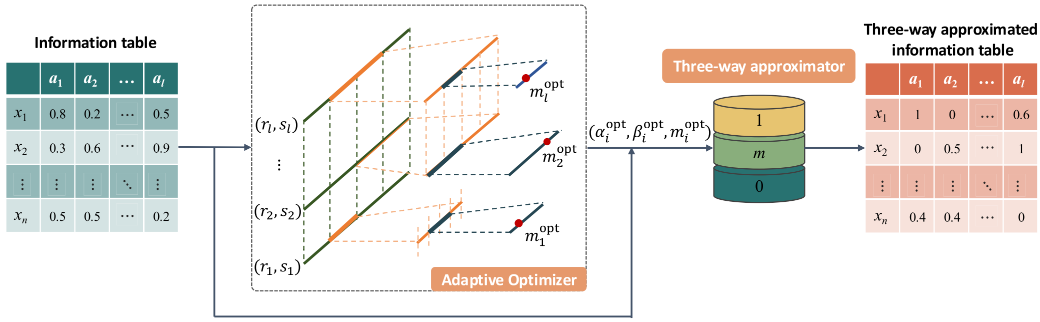

Fig. 5. Illustration of the asymmetric three-way approximation model.

TABLE VI
DATA INFORMATION

Datasets	Instances	Attributes	Classes	Feature types
BCWD	569	30	2	Real
WINE	178	13	3	Integer, Real
BTSC	748	4	2	Real
RCO	3810	7	2	Real
IONO	351	34	2	Integer, Real
IRIS	150	4	3	Real
PARK	197	22	2	Real
WDG1	5000	21	3	Real
ECOLI	336	7	8	Real
MHR	1013	6	3	Integer, Real

Note: The full names of the datasets are as follows: BCWD—Breast Cancer Wisconsin (Diagnostic) Dataset, WINE—Wine Dataset, BTSC—Blood Transfusion Service Center Dataset, RCO—Rice (Cammeo and Osmancik) Dataset, IONO—Ionosphere Dataset, IRIS—Iris Dataset, PARK—Parkinsons Dataset, WDG1—Waveform Database Generator (Version 1) Dataset, ECOLI—Ecoli Dataset, and MHR—Maternal Health Risk Dataset.

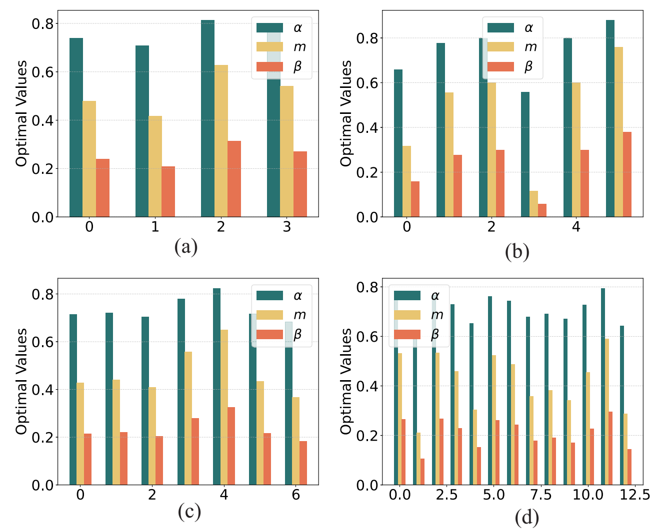

Fig. 6. Visualization of optimal thresholds for attributes across datasets. (a) IRIS. (b) MHR. (c) RCO. (d) WINE.

Tables II–V also show how α and β vary with λ_1, λ_2, μ_1, and μ_2. Specifically, α increases with λ_1 and decreases with λ_2, while β decreases with μ_1 and increases with μ_2. These trends are consistent with the results in Proposition 1.

2) Experiments on UCI Datasets: This section further analyzes the effect of the intermediate value on information loss through practical datasets. We obtained 10 datasets from the UC Irvine Machine Learning Repository [53], with details provided in Table VI. Each dataset is represented by an information table $\mathrm{IT} = (U, AT, f, V)$, where

1) U is the set of objects;
2) AT is the set of attributes;
3) f is the information function mapping each object to a value for each attribute in V;
4) $V = \cup_{a \in AT} V_a$, with $V_a \subseteq \mathbb{R}$ for each attribute a.

In our experiments, each attribute value space V_a is a subset of real numbers.

For each dataset, we compute its three-way approximation using the following steps (see Fig. 5).

1) *Compute near-optimal intermediate value:* For each attribute a_i, find the near-optimal intermediate value m_i^{opt} using an adaptive optimizer.
   a) Start with an initial search range $r_i = [0, 1]$ and step size $s_i = 0.1$. Find the optimal solution m_i^{opt} in the initial search.
   b) Update the search range $r_i = [m_i^{\mathrm{opt}} - s, m_i^{\mathrm{opt}} + s]$ (or $r = [m_i^{\mathrm{opt}}, m_i^{\mathrm{opt}} + s]$ or $r_i = [m_i^{\mathrm{opt}} - s, m_i^{\mathrm{opt}}]$ if m_i^{opt} is one of the endpoints of the search range) and reduce the step size to $s_i = \frac{m_i^{\mathrm{opt}}}{10}$. Find the optimal solution m_i^{opt} in this search.
   c) Repeat the process in the last step until the difference in information losses between two iterations is less than a threshold $\delta = 0.001$.
   d) Once m_i^{opt} is found, output the corresponding α_i^{opt} and β_i^{opt}.
2) *Approximate the information table:* Using the optimal values $(\alpha_i^{\mathrm{opt}}, \beta_i^{\mathrm{opt}}, m_i^{\mathrm{opt}})$, compute the new information function g for each object $x \in U$ and attribute $a_i \in AT$ as

$$g(x, a_i) = \begin{cases} 1, & f(x, a_i) \geq \alpha_i^{\mathrm{opt}} \\ m_i^{\mathrm{opt}}, & \beta_i^{\mathrm{opt}} < f(x, a_i) < \alpha_i^{\mathrm{opt}} \\ 0, & f(x, a_i) \leq \beta_i^{\mathrm{opt}}. \end{cases}$$

The resulting information table is $\mathrm{IT}^{3\mathrm{WA}} = (U, AT, g, V')$, where $V' = \{0, 1, m_1^{\mathrm{opt}}, m_2^{\mathrm{opt}}, \ldots, m_l^{\mathrm{opt}}\}$.

Algorithm 1 describes obtaining a three-way approximation of an information table. The first *for-loop* (lines 1–19) calculates the optimal values of m, α, and β for each attribute a_i. The third *for-loop* (lines 21–33) computes the three-valued information function using these optimal values. For n instances and l

Algorithm 1: Three-Way Approximation of Information Table.

input : Information table $\mathrm{IT} = (U, AT, f, V)$, loss factor $(\lambda_1, \lambda_2, \mu_1, \mu_2)$, adaptive factor δ.
output: Three-way approximated information table $\mathrm{IS}^{\mathrm{3WA}} = (U, AT, g, V')$.

1 **for** *each* $a_i \in AT$ **do**
2 $\quad s = 0.1$, $r = [0, 1]$, $m = 0$, $curr = 0.0$, $last = 1.0$
3 $\quad$ // $curr$ and $last$ are the current and last optimal information losses.
4 $\quad$ **while** $|last - curr| > \delta$ **do**
5 $\quad\quad last = curr$
6 $\quad\quad$ **for** $m \in r$ **do**
7 $\quad\quad\quad$ info-loss$[m]$ = information loss based on m
8 $\quad\quad\quad m = m + s$
9 $\quad\quad$ **end**
10 $\quad\quad$ // Compute information loss for each m in range r with step length s.
11 $\quad\quad curr = \min_m\{\text{info-loss}[m]\}$
12 $\quad\quad m^{\mathrm{opt}} = \arg\min_m\{\text{info-loss}[m]\}$
13 $\quad\quad r = [\max\{r_{\mathrm{left}}, m^{\mathrm{opt}} - s\}, \min\{r_{\mathrm{right}}, m^{\mathrm{opt}} + s\}]$
14 $\quad\quad$ // r_{left} and r_{right} are the left and right endpoints of the current r.
15 $\quad\quad s = \frac{r_{\mathrm{right}} - r_{\mathrm{left}}}{10}$
16 $\quad\quad m = r_{\mathrm{left}}$
17 $\quad$ **end**
18 $\quad m_i^{\mathrm{opt}} = m^{\mathrm{opt}}$, $\alpha_i^{\mathrm{opt}} = \frac{\lambda_1 + \lambda_2 m_i^{\mathrm{opt}}}{\lambda_1 + \lambda_2}$, $\beta_i^{\mathrm{opt}} = \frac{\mu_2 m_i^{\mathrm{opt}}}{\mu_1 + \mu_2}$
19 **end**
20 // Compute optimal thresholds for each attribute.
21 **for** $a_i \in AT$ **do**
22 $\quad$ **for** $x \in U$ **do**
23 $\quad\quad$ **if** $f(x, a_i) \geq \alpha_i^{opt}$ **then**
24 $\quad\quad\quad g(x, a_i) = 1$
25 $\quad\quad$ **endif**
26 $\quad\quad$ **if** $\beta_i^{opt} < f(x, a_i) < \alpha_i^{opt}$ **then**
27 $\quad\quad\quad g(x, a_i) = m_i^{\mathrm{opt}}$
28 $\quad\quad$ **endif**
29 $\quad\quad$ **if** $f(x, a_i) \leq \beta_i^{opt}$ **then**
30 $\quad\quad\quad g(x, a_i) = 0$
31 $\quad\quad$ **endif**
32 $\quad$ **end**
33 **end**
34 // Compute three-way approximation of f.
35 $V' = \{0, 1\} \cup \{m_i^{\mathrm{opt}}\}$

attributes, the computational complexity of the first *for-loop* is $O(l \times M \times 20)$, where M is the maximum number of *while-loop* iterations, and 20 is the maximum number of steps in the second *for-loop* (lines 6–9). The complexity of the third *for-loop* is $O(l \times n)$, making the overall complexity of Algorithm 1 $O(l \times M \times 20) + O(l \times n)$. Here, $O(l \times M \times 20)$ represents the complexity of finding the optimal values for m, α, and β, while $O(l \times n)$ represents the complexity of computing the three-valued information function.

The parameters $(\lambda_1, \lambda_2, \mu_1, \mu_2)$ are fixed at $(1, 1, 1, 1)$ for all datasets to determine the optimal value of m. Fig. 6 presents the optimal values of α, β, and m for datasets IRIS, MHR, RCO, and WINE. It demonstrates that the optimal values of m differ across attributes within the same dataset. For example, for the IRIS dataset, the optimal m values for each attribute are 0.48, 0.4167, 0.6271, and 0.5417. This reinforces the idea that setting m to a fixed value 0.5 is not always appropriate to minimize information loss.

B. Classification Performance Evaluation

In this section, we assess the classification performance of different models in three aspects: 1) a comparative analysis of classification accuracy; 2) a statistical test for performance significance; and 3) an evaluation of the impact of individual model parameters. We employed four classifiers—multinomial naive Bayes (MNB), logistic regression (LR), decision tree (DT), and feedforward neural network (FNN)—and tested them on ten benchmark datasets (details in Table VI). All numerical features were normalized to [0,1] using Min–Max normalization, and fivefold cross-validation was applied to all models.

1) Comparative Experiments: We compared the classification accuracy of the Baseline model (directly applying classification to normalized data), MMSS model [20], MESS model [21], S3WA model [47], [48], and three A3WA models. The A3WA-I model uses 0.5 as the intermediate value, A3WA-II takes the mean of uncertain membership values, and A3WA-III optimizes m in $0.1, 0.2, \ldots, 0.9$. All other factors are searched within $0.2, 0.4, \ldots, 1.0$.

Table VII summarizes each model's classification accuracy and standard deviation across the datasets for all classifiers. The experimental results demonstrate that the A3WA models consistently outperform the Baseline, MMSS, MESS, and S3WA models across most datasets and classifiers. Among the three A3WA variants, A3WA-III delivers the best overall performance with 23 highest accuracy instances across dataset-classifier combinations. A3WA-II follows with 11 cases, while A3WA-I achieves 6. The baseline records 5, MMSS 2, S3WA 1, and MESS none among the other models. A3WA-III achieves the highest accuracy on MNB (9 cases), LR (5 cases), and FNN (5 cases), highlighting its generalization capability, robustness, and adaptability across classification tasks and data complexities. In datasets like BCWD and IRIS, most models achieve similar high accuracy. However, the performance gap widens in datasets like RCO and ECOLI across MNB and FNN, where A3WA models—especially A3WA-III—show notable improvements. Overall, A3WA-III proves the most reliable, followed by A3WA-II, highlighting the effectiveness of adaptive parameter optimization in enhancing classification performance.

2) Statistical Test: We assessed the methods' significance using the Friedman test [54] with $\alpha = 0.05$ on 10 datasets. The null hypothesis assumes all methods perform equally, which is rejected if the p-value is below α. The Friedman test yielded p-values of 2.7557×10^{-8}, 3.4263×10^{-7}, 3.61×10^{-7}, and 9.4501×10^{-7} for the four classifiers, all significantly below 0.05, confirming statistical differences between the methods. To further analyze these differences, we applied Nemenyi's test [54], which compares the ranked performances of methods. The null hypothesis assumes no significant difference between methods and is rejected if the average rank difference exceeds the critical difference (CD). With $q_\alpha = 3.15$, $k = 7$, and $N = 10$, the calculated CD was 3.04.

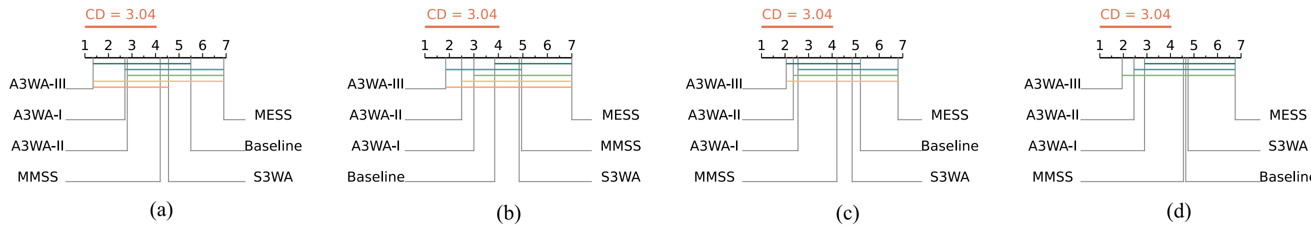

Fig. 7. Nemenyi test of different models across four classifiers. (a) MNB. (b) LR. (c) DT. (d) FNN.

TABLE VII
PERFORMANCE COMPARISON OF MODELS ACROSS CLASSIFIERS

Models	Dataset	MNB	LR	DT	FNN
Baseline	BCWD	83.47$_{\pm 2.45}$	97.01$_{\pm 0.91}$	93.32$_{\pm 0.69}$	94.03$_{\pm 1.02}$
MMSS		84.53$_{\pm 1.76}$	93.67$_{\pm 2.26}$	92.62$_{\pm 2.97}$	93.67$_{\pm 2.25}$
MESS		67.13$_{\pm 4.63}$	83.83$_{\pm 4.01}$	83.30$_{\pm 2.15}$	62.74$_{\pm 2.67}$
S3WA		91.74$_{\pm 1.44}$	96.31$_{\pm 1.40}$	94.37$_{\pm 1.64}$	95.25$_{\pm 0.91}$
A3WA-I		92.44$_{\pm 1.21}$	97.19$_{\pm 1.29}$	**97.19**$_{\pm 1.29}$	95.96$_{\pm 1.20}$
A3WA-II		87.87$_{\pm 1.21}$	95.96$_{\pm 1.89}$	93.85$_{\pm 2.08}$	95.95$_{\pm 1.21}$
A3WA-III		**92.44**$_{\pm 1.81}$	**97.54**$_{\pm 0.35}$	95.96$_{\pm 1.31}$	**96.14**$_{\pm 1.05}$
Baseline	WINE	95.51$_{\pm 2.23}$	**98.32**$_{\pm 1.37}$	89.90$_{\pm 4.14}$	96.05$_{\pm 2.90}$
MMSS		93.30$_{\pm 4.51}$	97.76$_{\pm 1.12}$	96.10$_{\pm 3.76}$	94.97$_{\pm 2.04}$
MESS		54.54$_{\pm 5.89}$	79.19$_{\pm 8.04}$	79.21$_{\pm 8.73}$	40.98$_{\pm 6.05}$
S3WA		93.81$_{\pm 1.19}$	95.49$_{\pm 2.29}$	92.70$_{\pm 2.90}$	94.37$_{\pm 1.83}$
A3WA-I		95.54$_{\pm 3.76}$	97.78$_{\pm 2.72}$	**97.78**$_{\pm 2.72}$	96.11$_{\pm 3.33}$
A3WA-II		93.87$_{\pm 4.77}$	98.30$_{\pm 1.39}$	96.67$_{\pm 4.08}$	**96.67**$_{\pm 3.24}$
A3WA-III		**96.10**$_{\pm 4.15}$	**98.32**$_{\pm 2.23}$	96.65$_{\pm 2.72}$	96.65$_{\pm 3.24}$
Baseline	BTSC	76.21$_{\pm 3.91}$	76.89$_{\pm 4.52}$	70.07$_{\pm 4.50}$	76.21$_{\pm 3.91}$
MMSS		76.62$_{\pm 4.53}$	76.89$_{\pm 4.48}$	78.09$_{\pm 4.55}$	76.21$_{\pm 3.91}$
MESS		76.21$_{\pm 3.91}$	76.48$_{\pm 4.15}$	76.08$_{\pm 4.06}$	76.21$_{\pm 3.91}$
S3WA		76.62$_{\pm 4.23}$	76.75$_{\pm 4.40}$	77.15$_{\pm 3.99}$	76.21$_{\pm 3.91}$
A3WA-I		76.88$_{\pm 4.26}$	77.02$_{\pm 4.14}$	78.48$_{\pm 3.32}$	**76.48**$_{\pm 4.36}$
A3WA-II		77.02$_{\pm 4.36}$	76.89$_{\pm 4.48}$	80.09$_{\pm 2.86}$	76.21$_{\pm 3.91}$
A3WA-III		**77.15**$_{\pm 4.21}$	**77.02**$_{\pm 4.58}$	**80.22**$_{\pm 2.80}$	76.21$_{\pm 3.91}$
Baseline	RCO	62.41$_{\pm 2.97}$	**92.86**$_{\pm 0.72}$	88.37$_{\pm 1.35}$	92.34$_{\pm 0.58}$
MMSS		72.28$_{\pm 1.54}$	90.76$_{\pm 0.99}$	90.87$_{\pm 0.76}$	90.60$_{\pm 0.54}$
MESS		57.22$_{\pm 1.74}$	60.79$_{\pm 2.14}$	60.60$_{\pm 2.14}$	59.08$_{\pm 2.18}$
S3WA		86.96$_{\pm 0.70}$	90.73$_{\pm 0.44}$	90.73$_{\pm 0.93}$	90.73$_{\pm 0.90}$
A3WA-I		87.51$_{\pm 1.45}$	90.79$_{\pm 0.47}$	91.00$_{\pm 0.52}$	90.58$_{\pm 0.71}$
A3WA-II		87.06$_{\pm 1.39}$	91.89$_{\pm 0.89}$	91.29$_{\pm 1.00}$	91.23$_{\pm 0.79}$
A3WA-III		**91.78**$_{\pm 1.03}$	92.70$_{\pm 0.72}$	**93.04**$_{\pm 0.64}$	**93.02**$_{\pm 0.61}$
Baseline	IONO	65.54$_{\pm 6.06}$	88.04$_{\pm 2.62}$	88.32$_{\pm 1.90}$	78.38$_{\pm 5.99}$
MMSS		**78.34**$_{\pm 4.59}$	90.04$_{\pm 2.35}$	92.60$_{\pm 2.45}$	88.03$_{\pm 1.50}$
MESS		64.11$_{\pm 4.81}$	82.35$_{\pm 5.07}$	80.91$_{\pm 3.06}$	78.36$_{\pm 4.84}$
S3WA		66.39$_{\pm 6.22}$	90.31$_{\pm 2.46}$	88.04$_{\pm 3.16}$	79.77$_{\pm 7.09}$
A3WA-I		68.68$_{\pm 7.15}$	91.46$_{\pm 2.35}$	92.31$_{\pm 1.70}$	83.77$_{\pm 4.07}$
A3WA-II		**78.34**$_{\pm 4.59}$	**91.74**$_{\pm 2.08}$	93.16$_{\pm 1.66}$	89.45$_{\pm 3.59}$
A3WA-III		75.21$_{\pm 3.61}$	91.46$_{\pm 2.35}$	**94.58**$_{\pm 1.67}$	**90.04**$_{\pm 3.66}$
Baseline	IRIS	63.33$_{\pm 2.98}$	92.67$_{\pm 3.27}$	96.00$_{\pm 3.27}$	91.33$_{\pm 5.81}$
MMSS		84.00$_{\pm 9.75}$	96.00$_{\pm 3.27}$	97.33$_{\pm 1.33}$	98.00$_{\pm 2.67}$
MESS		54.00$_{\pm 3.89}$	65.33$_{\pm 3.40}$	70.00$_{\pm 6.99}$	54.00$_{\pm 12.89}$
S3WA		70.67$_{\pm 5.73}$	**97.33**$_{\pm 2.49}$	97.33$_{\pm 2.49}$	95.33$_{\pm 3.40}$
A3WA-I		82.67$_{\pm 13.56}$	**97.33**$_{\pm 3.27}$	**98.00**$_{\pm 2.67}$	98.00$_{\pm 2.67}$
A3WA-II		84.00$_{\pm 9.75}$	**97.33**$_{\pm 2.49}$	**98.00**$_{\pm 2.67}$	98.00$_{\pm 2.67}$
A3WA-III		**90.67**$_{\pm 6.46}$	**97.33**$_{\pm 3.27}$	**98.00**$_{\pm 2.67}$	**98.67**$_{\pm 2.67}$
Baseline	PARK	85.13$_{\pm 6.15}$	84.10$_{\pm 6.57}$	82.05$_{\pm 8.73}$	82.05$_{\pm 6.88}$
MMSS		85.64$_{\pm 5.76}$	85.64$_{\pm 5.28}$	90.26$_{\pm 3.77}$	86.15$_{\pm 5.52}$
MESS		75.38$_{\pm 9.26}$	82.56$_{\pm 6.96}$	85.64$_{\pm 3.48}$	75.38$_{\pm 6.41}$
S3WA		82.05$_{\pm 5.13}$	86.15$_{\pm 4.76}$	89.74$_{\pm 3.97}$	84.62$_{\pm 3.97}$
A3WA-I		85.64$_{\pm 5.52}$	87.69$_{\pm 3.77}$	91.79$_{\pm 7.14}$	87.18$_{\pm 6.07}$
A3WA-II		85.64$_{\pm 5.76}$	86.15$_{\pm 5.52}$	**93.85**$_{\pm 4.17}$	**87.69**$_{\pm 4.10}$
A3WA-III		**86.67**$_{\pm 4.97}$	**88.72**$_{\pm 6.80}$	93.33$_{\pm 3.08}$	87.18$_{\pm 5.85}$
Baseline	WDG1	80.58$_{\pm 1.30}$	**87.02**$_{\pm 0.91}$	74.78$_{\pm 0.92}$	**85.78**$_{\pm 1.27}$
MMSS		79.90$_{\pm 1.15}$	81.72$_{\pm 0.84}$	72.18$_{\pm 0.66}$	81.42$_{\pm 1.18}$
MESS		35.24$_{\pm 3.13}$	39.52$_{\pm 1.30}$	38.80$_{\pm 1.24}$	35.60$_{\pm 2.78}$
S3WA		80.18$_{\pm 0.35}$	83.50$_{\pm 0.91}$	74.66$_{\pm 0.79}$	82.76$_{\pm 0.26}$
A3WA-I		80.60$_{\pm 1.07}$	84.08$_{\pm 0.74}$	**84.08**$_{\pm 0.74}$	84.12$_{\pm 0.84}$
A3WA-II		**80.86**$_{\pm 0.67}$	84.40$_{\pm 0.88}$	74.50$_{\pm 1.54}$	84.28$_{\pm 0.78}$
A3WA-III		80.60$_{\pm 1.07}$	84.36$_{\pm 1.27}$	74.98$_{\pm 0.55}$	83.94$_{\pm 0.77}$
Baseline	ECOlI	42.55$_{\pm 3.42}$	81.54$_{\pm 3.76}$	79.75$_{\pm 4.99}$	48.81$_{\pm 6.39}$
MMSS		46.12$_{\pm 4.28}$	76.79$_{\pm 3.31}$	**85.43**$_{\pm 1.64}$	54.44$_{\pm 12.21}$
MESS		42.55$_{\pm 3.42}$	45.23$_{\pm 3.59}$	52.97$_{\pm 1.62}$	42.55$_{\pm 3.42}$
S3WA		72.63$_{\pm 4.20}$	80.06$_{\pm 3.06}$	80.66$_{\pm 2.64}$	63.99$_{\pm 6.06}$
A3WA-I		72.63$_{\pm 4.20}$	81.85$_{\pm 2.51}$	82.45$_{\pm 3.62}$	66.08$_{\pm 5.68}$
A3WA-II		68.45$_{\pm 4.74}$	**83.33**$_{\pm 1.52}$	**85.43**$_{\pm 1.64}$	66.38$_{\pm 7.17}$
A3WA-III		**75.61**$_{\pm 2.65}$	82.74$_{\pm 3.18}$	83.94$_{\pm 2.67}$	**73.52**$_{\pm 5.90}$
Baseline	MHR	54.44$_{\pm 4.22}$	62.82$_{\pm 1.18}$	**82.05**$_{\pm 3.45}$	58.87$_{\pm 4.97}$
MMSS		59.27$_{\pm 3.80}$	66.96$_{\pm 2.14}$	70.12$_{\pm 3.89}$	64.10$_{\pm 3.39}$
MESS		44.28$_{\pm 3.45}$	52.86$_{\pm 4.55}$	55.32$_{\pm 5.20}$	51.87$_{\pm 4.13}$
S3WA		58.88$_{\pm 4.77}$	62.03$_{\pm 3.67}$	69.92$_{\pm 3.76}$	62.32$_{\pm 5.36}$
A3WA-I		61.14$_{\pm 3.90}$	66.86$_{\pm 3.22}$	71.20$_{\pm 3.90}$	65.97$_{\pm 2.88}$
A3WA-II		61.04$_{\pm 3.49}$	**67.75**$_{\pm 2.10}$	71.60$_{\pm 2.79}$	65.67$_{\pm 5.44}$
A3WA-III		**61.54**$_{\pm 3.77}$	67.45$_{\pm 4.63}$	71.20$_{\pm 3.90}$	**66.36**$_{\pm 4.73}$

Bold values indicate the highest classification accuracy for each dataset and classifier among all compared methods.

Table VIII presents the average ranks of each model, and Fig. 7 shows the test results. A3WA-III consistently ranks highest, outperforming other models across classifiers. It significantly differs from the Baseline and MESS on MNB, DT, and FNN. The

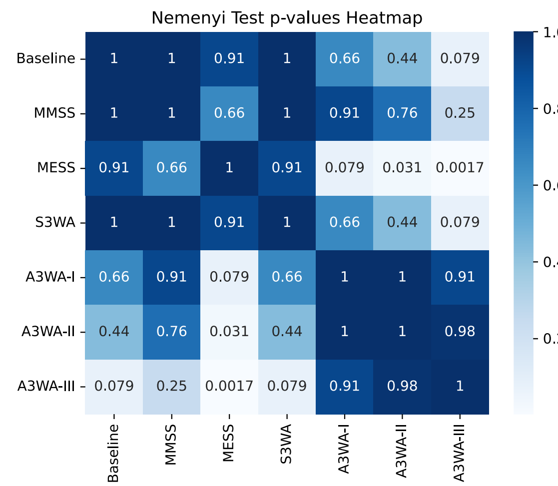

Fig. 8. Nemenyi test p-values heatmap.

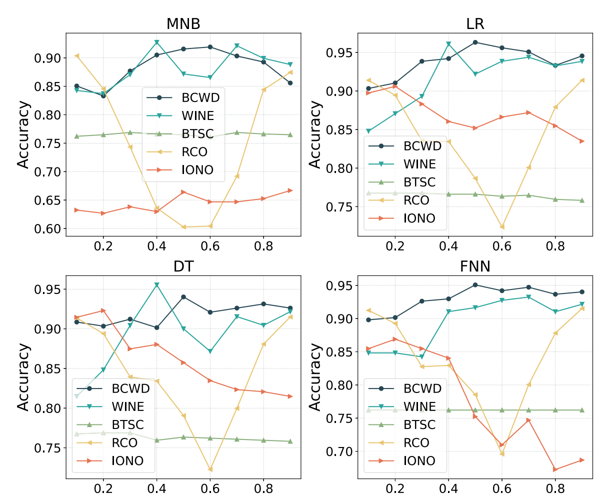

Fig. 9. Impact of the intermediate value m on classification accuracy across classifiers.

final row of Table VIII shows the overall average rank, indicating that a fixed intermediate value of 0.5 underperforms compared to the membership mean and globally optimized m (as seen in the A3WA variants), further emphasizing the inadequacy of the fixed-value approach. Fig. 8 visualizes the Nemenyi test p-value heatmap, confirming that A3WA-III outperforms baseline, MESS, and S3WA, with A3WA-II and A3WA-I performing similarly well. These results demonstrate the superior performance of A3WA models, with A3WA-III leading, followed by A3WA-II and A3WA-I.

Fig. 10. Impact of various factors on classification accuracy across classifiers. (a) Classification Accuracy with Variations in λ_1. (b) Classification Accuracy with Variations in λ_2. (c) Classification Accuracy with Variations in μ_1. (d) Classification Accuracy with Variations in μ_2.

TABLE VIII
AVERAGE RANKS OF MODELS ACROSS CLASSIFIERS

Classifiers	Baseline	MMSS	MESS	S3WA	A3WA-I	A3WA-II	A3WA-III
MNB	5.50	4.20	6.90	4.55	2.70	2.80	**1.35**
LR	3.85	4.95	7.00	4.85	3.00	2.50	**1.85**
DT	5.20	4.20	6.80	4.85	2.55	2.35	**2.05**
FNN	4.65	4.55	6.75	4.75	2.90	2.45	**1.95**
Averages	4.80	4.47	6.86	4.75	2.79	2.52	**1.80**

Bold values denote the lowest average rank among all methods, indicating the overall best performance.

3) Parameter Analysis on Classification Accuracy: We conducted five experiments to assess the impact of specific parameters (λ_1, λ_2, μ_1, μ_2, or m) on classification accuracy. In each experiment, the other four parameters were fixed, while the selected parameter was varied based on predefined values: $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \{0.2, 0.4, \ldots, 5.0\}$ and $m \in \{0.1, 0.2, \ldots, 0.9\}$. The fixed values for the parameters were $\lambda_1 = \lambda_2 = \mu_1 = \mu_2 = 1.0$ and $m = 0.5$.

We presented the results from five representative datasets. Fig. 9 shows the effect of the intermediate value m on classification accuracy across classifiers. The relationship between m and accuracy varies by dataset. For RCO, accuracy decreases as m increases, then rises again. For BCWD, accuracy improves initially, then decreases. For other datasets, accuracy fluctuates unpredictably without a clear trend. Fig. 10 illustrates the impact of λ_1, λ_2, μ_1, and μ_2 on classification accuracy, with each subplot containing four smaller subplots presenting results for different classifiers. The effect of each parameter differs across datasets and classifiers. For RCO, accuracy decreases with increasing λ_1 and μ_1, but increases with λ_2 and μ_2. For BCWD, λ_1 and λ_2 have minimal effect, while μ_1 reduces accuracy and μ_2 slightly improves it. For other datasets, the parameters show inconsistent or negligible effects across classifiers.

IV. CONCLUSION

This article presented an A3WA model with analytical solutions for determining optimal thresholds. A key finding was that when m was given, the optimal values of α and β were independent of the membership structure. However, the membership structure influenced the choice of m, even when α and β were fixed. Ablation experiments validated this distinction, further solidifying the model's theoretical framework.

Comparative experiments showed that our model flexibly delivered superior classification performance across various datasets and machine learning algorithms.

Looking forward, two key directions emerge for future work. First, since our model is built on information loss, exploring other A3WA models grounded in other uncertainty invariances will provide deeper insights and broaden their applicability. Second, applying our model in real-world scenarios is essential, with a particular emphasis on improving the interpretability of deep learning algorithms. This will ensure that the model enhances computational performance and supports clearer decision-making in practical applications. We aim to advance the theoretical development and practical utility of three-way approximation models through these future efforts.

References

[1] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, pp. 338–353, 1965.

[2] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, vol. 4. Hoboken, NJ, USA: Prentice Hall, 1995.

[3] H.-J. Zimmermann, *Fuzzy Set Theory–and Its Applications.* New York, NY, USA: Springer-Verlag, 2011.

[4] D. Wu and J. M. Mendel, "On the continuity of type-1 and interval type-2 fuzzy logic systems," *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 1, pp. 179–192, 2011.

[5] D. Wu, "On the fundamental differences between interval type-2 and type-1 fuzzy logic controllers," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 5, pp. 832–848, 2012.

[6] W. Pedrycz, "Shadowed sets: Representing and processing fuzzy sets," *IEEE Trans. Syst., Man, Cybernet., Part B*, vol. 28, no. 1, pp. 103–109, 1998.

[7] W. Pedrycz, "Shadowed sets: Bridging fuzzy and rough sets," in *Rough Fuzzy Hybridization. A New Trend in Decision-Making*. Singapore: Springer-Verlag, pp. 179–199, 1999.

[8] W. Pedrycz, "From fuzzy sets to shadowed sets: Interpretation and computing," *Int. J. Intell. Syst.*, vol. 24, no. 1, pp. 48–61, 2009.

[9] X. Deng and Y. Yao, "Decision-theoretic three-way approximations of fuzzy sets," *Inf. Sci.*, vol. 279, pp. 702–715, 2014.

[10] Y. Yao, S. Wang, and X. Deng, "Constructing shadowed sets and three-way approximations of fuzzy sets," *Inf. Sci.*, vol. 412–413, pp. 132–153, 2017.

[11] X. Zhao and Y. Yao, "Three-way fuzzy partitions defined by shadowed sets," *Inf. Sci.*, vol. 497, pp. 23–37, 2019.

[12] Y. Yao, "Set-theoretic models of three-way decision," *Granular Comput.*, vol. 6, no. 1, pp. 133–148, 2021.

[13] Y. Yao, "Symbols-meaning-value (SMV) space as a basis for a conceptual model of data science," *Int. J. Approx. Reasoning*, vol. 144, pp. 113–128, 2022.

[14] Y. Yao, "The DAO of three-way decision and three-world thinking," *Int. J. Approx. Reasoning*, vol. 162, 2023, Art. no. 109032.

[15] X. Zhao and B. Q. Hu, "Three-way decisions with decision-theoretic rough sets in multiset-valued information tables," *Inf. Sci.*, vol. 507, pp. 684–699, 2020.

[16] X. Zhao and D. Miao, "Isomorphic relationship between L-three-way concept lattices," *Cogn. Comput.*, vol. 14, pp. 1997–2019, 2022.

[17] Q. Zhang, Y. Chen, J. Yang, and G. Wang, "Fuzzy entropy: A more comprehensible perspective for interval shadowed sets of fuzzy sets," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 11, pp. 3008–3022, Dec. 2020.

[18] G. Cattaneo and D. Ciucci, "Shadowed sets and related algebraic structures," *Fundamenta Informaticae*, vol. 55, no. 3/4, pp. 255–284, 2003.

[19] G. Cattaneo and D. Ciucci, "Theoretical aspects of shadowed sets," in *Handbook of Granular Computing*, W. Pedrycz, A. Skowron, and V. Kreinovich, Eds., Hoboken, NJ, USA: Wiley, 2008, pp. 603–627.

[20] X. Deng and Y. Yao, "Mean-value-based decision-theoretic shadowed sets," in *Proc. Joint IFSA World Congr. NAFIPS Annu. Meeting (IFSA/NAFIPS)*, Edmonton, AB, Canada, Jun. 2013, pp. 1382–1387.

[21] M. Gao, Q. Zhang, F. Zhao, and G. Wang, "Mean-entropy-based shadowed sets: A novel three-way approximation of fuzzy sets," *Int. J. Approx. Reasoning*, vol. 120, pp. 102–124, 2020.

[22] J. Yang and Y. Yao, "Semantics of soft sets and three-way decision with soft sets," *Knowledge-Based Syst.*, vol. 194, 2020, Art. no. 105538.

[23] J. Yang and Y. Yao, "A three-way decision based construction of shadowed sets from Atanassov intuitionistic fuzzy sets," *Inf. Sci.*, vol. 577, pp. 1–21, 2021.

[24] H. Tahayori, A. Sadeghian, and W. Pedrycz, "Induction of shadowed sets based on the gradual grade of fuzziness," *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 5, pp. 937–949, Jun. 2013.

[25] Z. Luo, J. Hu, Q. Zhang, and G. Wang, "Induction of interval shadowed sets from the perspective of maintaining fuzziness," *Int. J. Approx. Reasoning*, vol. 153, pp. 219–238, 2023.

[26] J. Zhou, D. Miao, C. Gao, Z. Lai, and X. Yue, "Constrained three-way approximations of fuzzy sets: From the perspective of minimal distance," *Inf. Sci.*, vol. 502, pp. 247–267, 2019.

[27] Q. Zhang, D. Xia, K. Liu, and G. Wang, "A general model of decision-theoretic three-way approximations of fuzzy sets based on a heuristic algorithm," *Inf. Sci.*, vol. 507, pp. 522–539, 2020.

[28] M. A. Ibrahim and T. O. William-West, "A generalized cost-sensitive model for decision-theoretic three-way approximation of fuzzy sets," *Inf. Sci.*, vol. 570, pp. 638–667, 2021.

[29] J. Yang, X. Wang, G. Wang, Q. Zhang, N. Zheng, and D. Wu, "Fuzziness-based three-way decision with neighborhood rough sets under the framework of shadowed sets," *IEEE Trans. Fuzzy Syst.*, vol. 32, no. 9, pp. 4976–4988, Oct. 2024.

[30] Y. Zhang and J. Yao, "Determining strategies in game-theoretic shadowed sets," in *Proc. Inf. Process. Manage. Uncertainty Knowl.-Based Syst. Theory Found.*, 2018, pp. 736–747.

[31] Y. Zhang and J. Yao, "Game theoretic approach to shadowed sets: A three-way tradeoff perspective," *Inf. Sci.*, vol. 507, pp. 540–552, 2020.

[32] Q. Zhang, M. Gao, F. Zhao, and G. Wang, "Fuzzy-entropy-based game theoretic shadowed sets: A novel game perspective from uncertainty," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 3, pp. 597–609, Apr. 2022.

[33] M. Gao, Q. Zhang, F. Zhao, C. Wu, G. Wang, and D. Xia, "Constructing shadowed set based on game analysis of uncertainty and decision cost," *Appl. Soft Comput.*, vol. 147, 2023, Art. no. 110762.

[34] W. Pedrycz, "Interpretation of clusters in the framework of shadowed sets," *Pattern Recognit. Lett.*, vol. 26, no. 15, pp. 2439–2449, 2005.

[35] S. Mitra, W. Pedrycz, and B. Barman, "Shadowed c-means: Integrating fuzzy and rough clustering," *Pattern Recognit.*, vol. 43, no. 4, pp. 1282–1291, 2010.

[36] J. Zhou, W. Pedrycz, and D. Miao, "Shadowed sets in the characterization of rough-fuzzy clustering," *Pattern Recognit.*, vol. 44, no. 8, pp. 1738–1749, 2011.

[37] L. Chen, J. Zou, and C. P. Chen, "Image segmentation using shadowed c-means and kernel method," in *Proc. Int. Conf. Fuzzy Theory Appl*, 2013, pp. 374–379.

[38] J. Zhou, Z. Lai, C. Gao, D. Miao, and X. Yue, "Rough possibilistic c-means clustering based on multigranulation approximation regions and shadowed sets," *Knowl.-Based Syst.*, vol. 160, pp. 144–166, 2018.

[39] H. Zhang, T. Zhang, W. Pedrycz, C. Zhao, and D. Miao, "Improved adaptive image retrieval with the use of shadowed sets," *Pattern Recognit.*, vol. 90, pp. 390–403, 2019.

[40] T. O. William-West and M. A. Ibrahim, "On shadowed set approximation methods," *Soft Comput.*, vol. 27, no. 8, pp. 4463–4482, 2023.

[41] S. He, X. Pan, and Y. Wang, "A shadowed set-based TODIM method and its application to large-scale group decision making," *Inf. Sci.*, vol. 544, pp. 135–154, 2021.

[42] C. Wu, Q. Zhang, F. Zhao, Y. Cheng, and G. Wang, "Three-way recommendation model based on shadowed set with uncertainty invariance," *Int. J. Approx. Reasoning*, vol. 135, pp. 53–70, 2021.

[43] L. Zhang, Y. Yao, and P. Zhu, "Shadowed set approximations of L-fuzzy sets," *Inf. Sci.*, vol. 679, 2024, Art. no. 121094.

[44] Z. Wang et al., "M-MSSEU: Source-free domain adaptation for multi-modal stroke lesion segmentation using shadowed sets and evidential uncertainty," *Health Inf. Sci. Syst.*, vol. 11, no. 1, 2023, Art. no. 46.

[45] A. Bose and K. Mali, "Gradual representation of shadowed set for clustering gene expression data," *Appl. Soft Comput.*, vol. 83, 2019, Art. no. 105614.

[46] K. Cai, H. Zhang, W. Pedrycz, and D. Miao, "SSS-Net: A shadowed-sets-based semi-supervised sample selection network for classification on noise labeled images," *Knowl.-Based Syst.*, 2023, Art. no. 110732.

[47] X. Yue, J. Zhou, Y. Yao, and D. Miao, "Shadowed neighborhoods based on fuzzy rough transformation for three-way classification," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 5, pp. 978–991, Jun. 2020.

[48] X. Yue, S. Liu, Q. Qian, D. Miao, and C. Gao, "Semi-supervised shadowed sets for three-way classification on partial labeled data," *Inf. Sci.*, vol. 607, pp. 1372–1390, 2022.

[49] Y. Zhang, T. Zhang, C. Peng, F. Ma, and W. Pedrycz, "Rough fuzzy k-means clustering based on parametric decision-theoretic shadowed set with three-way approximation," *Int. J. Fuzzy Syst.*, pp. 1–18, 2024.

[50] C. Jiang, Z. Li, and J. Yao, "A shadowed set-based three-way clustering ensemble approach," *Int. J. Mach. Learn. Cybern.*, vol. 13, no. 9, pp. 2545–2558, 2022.

[51] T. William-West, A. F. D. Kana, and M. A. Ibrahim, "Shadowed-set-based three-way clustering methods: An investigation of new optimization-based principles," *Inf. Sci.*, vol. 591, pp. 1–24, 2022.

[52] J. Zhou, Z. Lai, D. Miao, C. Gao, and X. Yue, "Multigranulation rough-fuzzy clustering based on shadowed sets," *Inf. Sci.*, vol. 507, pp. 553–573, 2020.

[53] D. Dua and C. Graff, "UCI machine learning repository," 2017. [Online]. Available: http://archive.ics.uci.edu/ml

[54] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *J. Mach. Learn. Res.*, vol. 7, pp. 1–30, 2006.

Xuerong Zhao received the B.Sc. degree in information and computational science from the University of Jinan, Jinan, China, in 2006, and the M.Sc. and Ph.D. degrees in computational mathematics from Wuhan University, Wuhan, China, in 2012 and 2015, respectively.

She is currently an Associate Professor with the College of Information, Mechanical, and Electrical Engineering, Shanghai Normal University, Shanghai, China. She completed a Postdoctoral Fellowship with Tongji University, Shanghai. She has an impressive publication record in premier journals, including *Information Sciences, Knowledge-Based Systems*, etc. Her research interests include rough sets, three-way decision, formal concept analysis, data mining, granular computing, computer vision, and more.

Duoqian Miao received the B.Sc. degree in fundamental mathematics from Shanxi University, Shanxi, China, in 1985, the M.Sc. degree in probability and statistics from Shanxi University, Shanxi, China, in 1991, and the Ph.D. degree in pattern recognition and intelligent systems from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 1997.

He is currently a Professor with the School of Electronics and Information Engineering, Tongji University, Shanghai, China. His work continues to advance intelligent systems and computational methods in artificial intelligence. His research interests include rough sets, soft computing, machine learning, and intelligent systems.

Prof. Miao is a Fellow of the International Rough Set Society and a Member of the Chinese Association for Artificial Intelligence (CAAI). He has published over 180 papers in leading journals, such as IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Pattern Analysis and Machine Intelligence, *Pattern Recognition*, and *Information Sciences*, as well as top-tier international conferences like AAAI and CVPR. He holds an h-index of 39, reflecting his significant contributions to the academic community. He is also an Associate Editor for Information Sciences, CAAI Transactions on Intelligence Technology, and the International Journal of Approximate Reasoning.

Yiyu Yao received the B.E. degree in computer science from Xi'an Jiaotong University, Xi'an, China, in 1983, and the M.Sc. and Ph.D. degrees in computer science from the University of Regina, Regina, SK, Canada, in 1988 and 1991, respectively.

He is currently a Professor of computer science with the University of Regina. He has authored or coauthored over 400 papers and was recognized as a Highly Cited Researcher from 2015 to 2019. His research interests include three-way decision, granular computing, rough sets, formal concept analysis, information retrieval, data mining, machine learning, and web intelligence. He proposed the theory of three-way decision, the decision-theoretic rough set model, and the triarchic theory of granular computing. He is also an Associate Editor for *Information Sciences* and the *International Journal of Approximate Reasoning*.

Witold Pedrycz (Life Fellow, IEEE) received the M.Sc. degree in computer science in 1977 and the Ph.D. degree in computer engineering in 1980 from the Silesian University of Technology, Gliwice, Poland, and the D.Sc. (habilitation) degree in systems science from the Polish Academy of Sciences, Poland, in 1984.

He is currently a Professor with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada. He is also with the Systems Research Institute of the Polish Academy of Sciences, Warsaw, Poland. His main research interests include computational intelligence, granular computing, and machine learning.

Dr. Pedrycz is a Foreign Member of the Polish Academy of Sciences and a Fellow of the Royal Society of Canada. He was the recipient of several awards including the Norbert Wiener Award from the IEEE Systems, Man, and Cybernetics Society, IEEE Canada Computer Engineering Medal, a Cajastur Prize for Soft Computing from the European Centre for Soft Computing, a Killam Prize, a Fuzzy Pioneer Award from the IEEE Computational Intelligence Society, and 2019 Meritorious Service Award from the IEEE Systems Man and Cybernetics Society. He serves as the Editor-in-Chief of *WIREs Data Mining and Knowledge Discovery* (Wiley) and Coeditor-in-Chief of *Journal of Data Information and Management* (Springer).