Shadowed Neighborhoods Based on Fuzzy Rough Transformation for Three-Way Classification

Xiaodong Yue, *Member, IEEE*, Jie Zhou, *Member, IEEE*, Yiyu Yao, and Duoqian Miao

***Abstract*—Neighborhoods form a set-level approximation of data distribution for learning tasks. Due to the advantages of data generalization and nonparametric property, neighborhood models have been widely used for data classification. However, the existing neighborhood-based classification methods rigidly assign a certain class label to each data instance and lack the strategies to handle the uncertain instances. The far-fetched certain classification of uncertain instances may suffer serious risks. To tackle this problem, in this article, we propose a novel shadowed set to construct shadowed neighborhoods for uncertain data classification. For the fuzzy–rough transformation in the proposed shadowed set, a step function is utilized to map fuzzy neighborhood memberships to the set of triple typical values $\{0, 1, 0.5\}$ and thereby partition a neighborhood into certain regions and uncertain boundary (neighborhood shadow). The threshold parameter in the step function for constructing shadowed neighborhoods is optimized through minimizing the membership loss in the mapping of shadowed sets. Based on the constructed shadowed neighborhoods, we implement a three-way classification algorithm to distinguish data instances into certain classes and uncertain case. Experiments validate the proposed three-way classification method with shadowed neighborhoods is effective in handling uncertain data and reducing the classification risk.**

***Index Terms*—Fuzzy rough transformation, shadowed neighborhood, three-way classification, uncertain data analysis.**

I. Introduction

NEIGHBORHOODS are constructed through grouping neighboring data instances into sets [1]. In contrast to K-Nearest Neighbors as instance prototypes [2]–[4], neighborhoods provide the set-level prototypes and thus facilitate the data generalization [5], [6]. Moreover, neighborhood models are generally nonparametric and need not assume the probability distribution of data, which make the neighborhood-based learning easy to implement and flexible to data diversity [7], [8]. The union of the homogeneous neighborhoods belonging to the same class approximates the data distribution for classification [9], [10]. The classifications based on neighborhoods were proven to be more efficient than the classifications based on nearest-neighbor search [11].

However, the existing neighborhood-based classification methods rigidly assign a certain class label to each data instance and lack the strategies to handle the instances with uncertainty. The methodology of uncertain data classification is very helpful in reducing the decision risk and in the meantime improves the decision efficiency through human–machine cooperation, and therefore plays an important role in decision support systems [12]. For an example, when we apply the neighborhood-based classification methods to implement a computer-aided diagnosis (CAD) system for liver cancer, it is required to classify the uncertain tumors for further cautious diagnosis and certain far-fetched classifications produced by the system may cause serious costs [13].

Aiming to tackle the limitation of neighborhood models for uncertain data classification, in this article, we utilize shadowed sets [14] to extend the traditional neighborhoods to shadowed ones and thereby propose a three-way classification method based on the shadowed neighborhoods. To integrate the two important paradigms of granular computing [15], [16]: Rough sets [17], [18] and fuzzy sets [19], [20], fuzzy rough sets [21], [22] have been widely investigated to achieve the unified methodology for uncertain data analysis [23]–[25]. Based on the fuzzy–rough transformation, shadowed sets are constructed through mapping fuzzy memberships into a triplet set $\{0, [0, 1], 1\}$ [26]. With the triple elements of shadowed sets, a fuzzy concept is tri-partitioned to form a rough representation which consists of certain positive region (denoted by 1), certain negative region (denoted by 0), and uncertain shadow region (denoted by $[0, 1]$). The traditional shadowed sets balance the uncertainty variations on certain and uncertain regions [26], which facilitate the uncertain data clustering [27] but may not suit supervised learning tasks. Motivated by this, we propose a novel shadowed set on fuzzy neighborhood memberships to construct the shadowed neighborhoods of certain regions and uncertain boundary (neighborhood shadow) to classify uncertain data.

To implement the uncertain classification based on shadowed neighborhoods, we refer to the methodology of three-way decisions (3WD) [28], [29] to design a three-way classification strategy. In the process of three-way decision making, decision rules are generated through tripartitioning data space into positive, negative, and boundary regions. Like the union of

Manuscript received May 15, 2019; revised January 1, 2020 and February 14, 2020; accepted February 18, 2020. Date of publication March 9, 2020; date of current version May 4, 2020. This work was supported in part by National Natural Science Foundation of China under Grant 61976134, Grant 61573235, Grant 61991410, Grant 61991415 and in part by Open Project Foundation of Intelligent Information Processing Key Laboratory of Shanxi Province Grant CICIP2018001. *(Corresponding author: Xiaodong Yue.)*

Xiaodong Yue is with the School of Computer Engineering and Science, Shanghai University, Shanghai 200444, China (e-mail: yswantfly@shu.edu.cn).

Jie Zhou is with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China (e-mail: jie_jpu@163.com).

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

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

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TFUZZ.2020.2979365

neighborhoods forms an approximation of data distribution for classification, the union of the shadowed neighborhoods forms a tripartitioned approximation of data distribution for three-way classification. The data instances will be classified into a certain class or uncertain case according to their locations with respect to the shadowed neighborhoods, such as the positive regions of the neighborhoods of same class certainly determine the class of instances but the neighborhood shadows have uncertainty for classification. The contributions of this article are summarized as follows.

1) *Construct and Optimize Shadowed Neighborhoods for Modeling Uncertain Data.*
   We propose a novel shadowed set on fuzzy neighborhood memberships to construct shadowed neighborhoods. In the proposed shadowed set, a step function is utilized to map neighborhood memberships to the set of triple typical values $\{0, 1, 0.5\}$ and thereby partitions a neighborhood into the certain positive region, negative region, and uncertain boundary region. Through minimizing the information loss in the transformation from fuzzy memberships to the shadowed set, we obtain the optimum threshold in the step function to optimize the construction of shadowed neighborhoods.
2) *Implement a Three-way Classification Algorithm With Shadowed Neighborhoods (3WC-SNB).*
   Based on the approximation of global data distribution formed by the shadowed neighborhoods, we design a group of three-way classification rules for both the data instances within and beyond neighborhoods, and also implement a 3WC-SNB to distinguish data instances into certain classes and uncertain case.

The rest of this article is organized as follows. Section II briefly introduces the preliminaries of shadowed sets and three-way decisions. Section III introduces the shadowed neighborhood model, which includes neighborhood membership formulation, shadowed neighborhood construction and optimization. Section IV presents a 3WC-SNB. In Section V, experimental results validate the effectiveness of the proposed method for uncertain data classification. Finally, Section VI concludes the article.

II. Preliminaries

A. Shadowed Sets of Fuzzy–Rough Transformation

As fuzzy rough sets [21], [22], shadowed sets [14], [26] were proposed by Pedrycz to bridge rough sets [17], [18] and fuzzy sets [19], [20] and thereby provide an effective tool to model and analyze the concepts with uncertainty. Shadowed sets are constructed through the fuzzy–rough transformation of fuzzy sets. In the fuzzy–rough transformation, the fuzzy memberships $\mu_A(x)$ of data instances $x \in X$ are mapped into a triplet set $\{0, [0, 1], 1\}$ and the mapping is formulated as $S^{\alpha}_{\mu_A} : X \to \{0, [0, 1], 1\}$. Referring to the fuzzy–rough sets [30], [31], the values 0 and 1 denote the certain negative region and certain positive region, and the interval [0, 1] denotes the uncertain region.

In the mapping of shadowed sets $S^{\alpha}_{\mu_A}$, $\alpha \in [0, 0.5]$ is the threshold parameter to tripartition the fuzzy memberships as

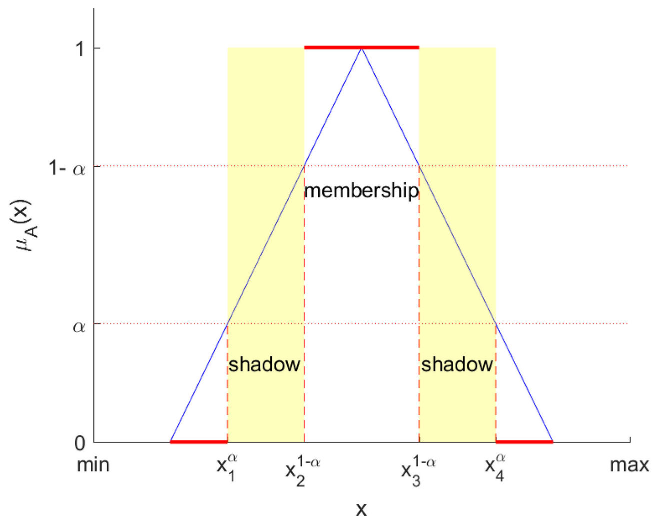

Fig. 1. Shadowed set of triangular membership function.

$$S^{\alpha}_{\mu_A}(x) = \begin{cases} 1, & \mu_A(x) \geq 1-\alpha \\ [0,1], & \alpha < \mu_A(x) < 1-\alpha \\ 0, & \mu_A(x) \leq \alpha. \end{cases} \quad (1)$$

The tripartition of fuzzy memberships forms a shadowed concept representation. The low memberships of instances no more than α will be reduced to the certain negative membership 0, the high memberships no less than $1-\alpha$ will be elevated to the certain positive membership 1, and the uncertain instances whose memberships locating in the interval $(\alpha, 1-\alpha)$ constitute the shadow area. The uncertainty of a shadowed set is measured by the number of the uncertain instances in the shadowed area.

Fig. 1 illustrates a shadowed set constructed on a triangular membership function. It can be found that the transformation from fuzzy memberships to a shadowed set relocates the uncertainty. The uncertainty in the positive and negative regions is reduced, and in the meantime, the uncertainty in the shadowed area is increased. Based on this, Pedrycz established the objective of uncertainty invariance to optimize the threshold parameter to construct shadowed sets.

Given a fuzzy membership function μ_A, for any data instance $x_i \in X$, its membership $\mu_A(x_i)$ is briefly denoted as μ_i. The uncertainty variance of transforming fuzzy memberships into a shadowed set [14], [15] is formulated as

$$V(\alpha) = \left| \sum_{\mu_i \leq \alpha} \mu_i + \sum_{\mu_i \geq 1-\alpha} (1-\mu_i) - card\{x_i \in X | \alpha < \mu_i < 1-\alpha\} \right|. \quad (2)$$

The uncertainty variance $V(\alpha)$ consists of two parts, namely, the uncertainty decrement of membership loss in the certain regions and the uncertainty increment in the uncertain region, which is represented by the number of uncertain instances in the shadow. Besides the membership loss, we can also interpret the uncertainty variance from the view of the areas of memberships [32], [33]

$$V(\alpha) = |\text{ElevatedArea}(S^{\alpha}_{\mu_A}) + \text{ReducedArea}(S^{\alpha}_{\mu_A}) - \text{ShadowArea}(S^{\alpha}_{\mu_A})|. \quad (3)$$

The optimum threshold parameter α^* should balance the shadowed area and the changing areas of memberships, i.e., the tradeoff between uncertainty and membership loss. $\alpha^* = \arg\min_{\alpha} V(\alpha)$, $V(\alpha) = 0$ will lead to the optimum membership threshold α^*.

Pedrycz's shadowed sets have been investigated and extended. Yao summarized the optimization strategies to construct shadowed sets in the framework of three-way decision theory, which include the strategies for minimizing distance and achieving the least cost [32]. Tahayori constructed the shadowed sets based on a gradual grade of fuzziness [34]. Nguyen proposed a distance-based shadowed approximation method to transform fuzzy recommendations to determined ones [35]. Grzeforzewski presented a shadowed set approximation to simplify fuzzy numbers, which also provided the interval and trapezoidal approximation methods for fuzzy sets [36]. Zhang proposed the game-theoretic shadowed sets, in which the thresholds of three-way approximation were determined by the principle of tradeoff with games [37].

Besides the construction of shadowed sets, they have been widely used to implement soft clusterings of data with uncertainty. Through mapping the fuzzy cluster memberships to a shadowed set with tripartition structure, fuzzy clustering [38], [39], and rough clustering [40], [41] can be represented in a uniform framework of shadowed clustering [27]. Based on this, the optimization strategies for constructing shadowed sets can be also utilized to optimize the threshold parameters of fuzzy and rough clusterings. Mitra proposed a shadowed C-means algorithm which integrates fuzzy and rough clustering [42]. And the rough–fuzzy clustering methods were also reinvestigated from the view of shadowed sets [43]. Zhou proposed a rough fuzzy clustering method based on shadowed sets, in which the clusters containing uncertain instances are modeled by shadowed sets and the thresholds for partitioning the certain and uncertain regions of clusters are determined through optimizing the shadowed sets [44], [45]. In general, the existing shadowed sets aim to maintain data uncertainty and the research focuses on the concept approximation and the applications of shadowed sets for uncertain data clustering. For the supervised learning tasks, such as data classification and regression, the related works are very limited.

B. Methodologies of Three-Way Decisions

Many soft computing models for leaning uncertain concepts, such as interval sets, many-valued logic, rough sets, fuzzy sets, and shadowed sets, have the common property of tripartitioning [28], [46]. Motivated by this, the methodology of 3WD is proposed as as an extension of the commonly used binary-decision model through adding a third option [29]. In general, the approach of 3WD divides the universe into the positive, negative, and boundary regions which denote the regions of acceptance, rejection, and noncommitment for ternary classifications. Specifically, for data classification, if the data instances partially satisfy the classification criteria, it is difficult to directly identify them without uncertainty. Instead of making a binary decision, we use thresholds on the degrees of satisfiability to make one of three decisions, i.e., accept, reject, or noncommitment. The third option may also be referred to as a deferment decision that requires further judgments.

With the ordered evaluation of acceptance, the three regions of decisions are formally defined through thresholding the evaluation values. Suppose $(L, \preceq)$ is a totally ordered set of evaluation values, in which $\preceq$ is a total order. For two thresholds $\alpha \prec \beta$, suppose the set of the values for acceptance is given by $L^+ = \{t \in L | t \succeq \alpha\}$ and the set for rejection is $L^- = \{b \in L | b \preceq \beta\}$. For an evaluation function $v : U \rightarrow L$, the positive, negative, and boundary regions are defined as

$$\begin{aligned} \text{POS}_{\alpha,\beta}(v) &= \{x \in U | v(x) \succeq \alpha\} \\ \text{NEG}_{\alpha,\beta}(v) &= \{x \in U | v(x) \preceq \beta\} \\ \text{BND}_{\alpha,\beta}(v) &= \{x \in U | \alpha \prec v(x) \prec \beta\}. \end{aligned} \quad (4)$$

Various kinds of decision-making methods have been reinvestigated within the framework of 3WD [47]–[49]. Three-way decision models were established from the perspectives of fuzzy sets, hesitant fuzzy sets, and interval-valued sets, respectively [50]–[52]. The three-way decision model was also revisited and extended from the views of game theory [53], sequential decision making [54], and formal concept analysis [55]. Besides, 3WD were utilized to construct the methods of uncertain clustering [56], [57], cost-sensitive classification [58], [59], and dynamic data classification [60]. Through integrating with machine learning methods, three-way decisions have been widely applied in the fields of recommendation system [61], network security [62], management analysis [63], social networks [64], natural language processing [65], disease diagnosis [13], and software detection [66]. Referring to the methodology of 3WD, we expect to reformulate neighborhoods with shadowed sets and thereby implement a 3WD method for uncertain data analysis.

III. Shadowed Neighborhoods

A. Fuzzy Neighborhood Membership

To construct the shadowed neighborhoods for classification, first we construct certain neighborhoods for data classification and fuzzify the neighborhoods to formulate the fuzzy neighborhood memberships. For a data instance x, its neighborhood consists of the surrounding instances with the same class.

Definition 1. Neighborhood [9]: Given a data instance $x \in X$, the neighborhood $O(x)$ of x is defined as

$$O(x) = \{y \mid d(x, y) \leq \eta, y \in X\}, \quad (5)$$

where $d(x, y)$ is the distance between the instances x and y, η denotes the radius of the neighborhood.

To handle the mixed-type data of both numerical and symbolic attributes, we adopt HEOM (Heterogeneous Euclidean-Overlap

Metric) function [67] as the distance measure to construct neighborhoods. To guarantee all the instances in the neighborhood belonging to the same class, i.e., the neighborhood homogeneity, we adopt the measures of nearest hit $NH(x)$ and nearest miss $NM(x)$ of the neighborhood center x to calculate the neighborhood radius referring to the strategy of neighborhood construction in [68]. $NH(x)$ is defined as the nearest instance to x with the same class label and $NM(x)$ is the nearest instance to x, which belongs to different classes. The neighborhood radius is calculated by $\eta = d(x, NM(x)) - 0.01 \times d(x, NH(x))$. Obviously, all the instances within the neighborhood of radius η belong to the same class as x.

The union of all the neighborhoods forms a covering of data, in which some neighborhoods may be contained in others, thus we further remove the redundant neighborhoods to simplify the model [69]. The remaining neighborhoods actually provide an approximation of global data distribution on set level and the instances within neighborhoods are uniformly distributed. Next we formulate the membership distribution of neighborhoods according to the distances from instances to neighborhood centers.

Definition 2. Neighborhood Membership: Given an instance x and a neighborhood $O(x_k)$, x_k is the neighborhood center, the membership of x belonging to $O(x_k)$ is defined based on the distance between x and x_k

$$\mu_{O(x_k)}(x) = 1 - \frac{1}{1 + e^{-t[d(x,x_k)-\eta]}} = \frac{e^{-t[d(x,x_k)-\eta]}}{1 + e^{-t[d(x,x_k)-\eta]}}. \tag{6}$$

The formula of neighborhood membership is a logistic function of "S" shape, in which $d(x, x_k)$ is the distance between x and x_k, $t \geq 1$ is the function order, and the neighborhood radius $\eta > 0$ is adopted as the function bias.

The neighborhood membership $\mu_{O(x_k)}(x) \in (0, 1)$. It can be found that, for the instance locating at the neighborhood boundary, i.e., $d(x, x_k) = \eta$, its neighborhood membership $\mu_{O(x_k)}(x) = 0.5$ and the membership decreases as the distance between data instance and neighborhood center increasing. In the next paragraphs, we briefly denote $\mu_{O(x_k)}(x)$ as $\mu_k(x)$.

B. Shadowed Neighborhood Construction

Based on the fuzzy–rough transformation of shadowed sets, we can transform the fuzzy neighborhood memberships of instances into rough ones and formulate a shadowed representation of neighborhoods. Different from the traditional shadowed sets mapping fuzzy memberships to $\{0, 1, [0, 1]\}$ as introduced in Section II, we propose a novel shadowed set which utilizes a step function to map fuzzy neighborhood memberships to the set of triple values $\{0, 1, 0.5\}$ for uncertain data classification. Specifically, the low memberships no more than α will be further reduced to 0 and the high memberships no less than $1 - \alpha$ will be elevated to 1, and the most uncertain membership value "0.5" is adopted to unify the neighborhood memberships of all the uncertain instances in the interval $(\alpha, 1 - \alpha)$. The shadowed neighborhood based on the shadowed set is defined as follows.

Definition 3. Shadowed Neighborhood: Given a neighborhood membership $\mu_k(x)$ and a threshold $\alpha \in [0, 0.5]$, the shadowed neighborhood is constructed through defining a shadowed

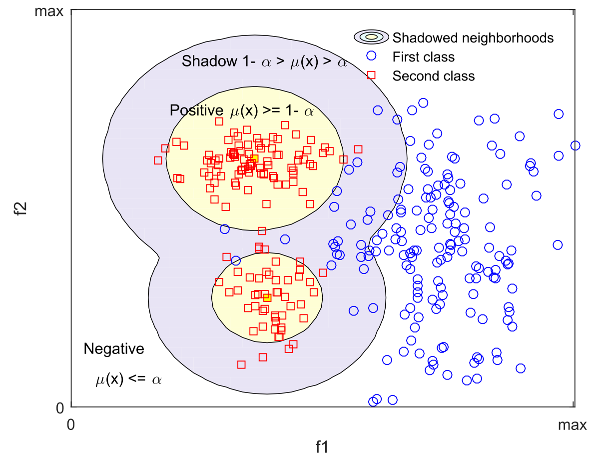

Fig. 2. Shadowed neighborhoods for binary classification.

set mapping of the neighborhood membership as

$$N_{\mu_k}^{\alpha}(x) = \begin{cases} 1, & \mu_k(x) \geq 1 - \alpha \\ 0.5, & \alpha < \mu_k(x) < 1 - \alpha \\ 0, & \mu_k(x) \leq \alpha. \end{cases} \tag{7}$$

The mapping of shadowed neighborhood $N_{\mu_k}^{\alpha}(x)$ utilizes a step function to approximate the neighborhood membership $\mu_k(x)$ and partitions the space into three regions according to the neighborhood belongingness, i.e., the *positive region* represented by membership grade 1, the *negative region* represented by membership grade 0, and the *boundary region* represented by membership grade 0.5, which forms the *neighborhood shadow*. For the three regions of a shadowed neighborhood, the positive region represents the data instances which certainly belong to the neighborhood, the negative region represents the instances which are certainly beyond the neighborhood, and the boundary region (neighborhood shadow) consists of the instances which are uncertain to belong to the neighborhood. Fig. 2 shows the shadowed neighborhoods of the data instances of one class for binary classification.

From the formula (7), we know that a shadowed neighborhood is constructed through discretizing quantitative neighborhood memberships using a step function to obtain qualified representations of neighborhood belongingness. The memberships of the instances in the positive region are elevated from $[1 - \alpha, 1]$ to 1, the memberships in the negative region are reduced from $[0, \alpha]$ to 0, and in the boundary region, the memberships ranging in $(\alpha, 1 - \alpha)$ are simplified to a unified value 0.5. The transformation from neighborhood membership $\mu_k(x)$ to a shadowed set $N_{\mu_k}^{\alpha}(x)$ causes the *membership loss* which is formulated as

$$L(\alpha) = \lambda \cdot \left[\sum_{\mu_k(x) \leq \alpha} \mu_k(x) + \sum_{\mu_k(x) \geq 1-\alpha} (1 - \mu_k(x)) \right] + \sum_{\alpha < \mu_k(x) < 1-\alpha} |0.5 - \mu_k(x)|. \tag{8}$$

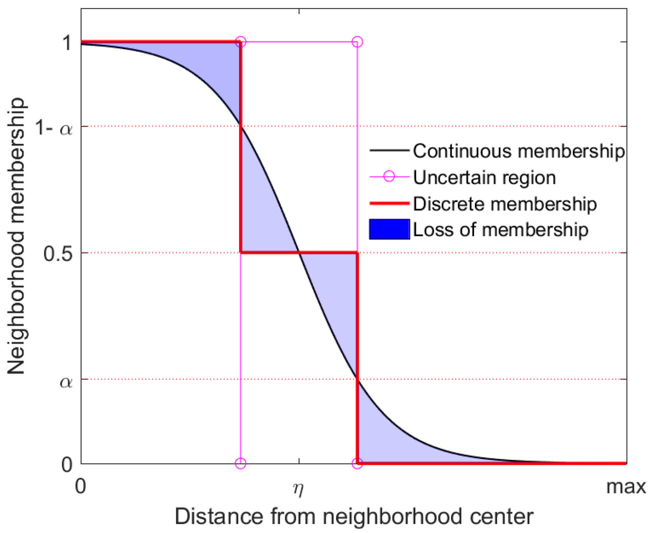

Fig. 3. Transformation from neighborhood membership to shadowed set.

$L(\alpha)$ consists of the membership losses in the certain positive region, negative region, and uncertain boundary region respectively. $\lambda > 0$ is the factor to balance the membership loss of the certain regions and uncertain region and we set $\lambda = 0.1$ as default. Fig. 3 illustrates the transformation from the neighborhood membership to a shadowed set and the corresponding membership loss. We find that for a given membership function (or a set of memberships), the membership loss is determined by the threshold α, thus we can optimize the threshold to construct shadowed neighborhoods through minimizing the membership loss.

C. Optimization of Shadowed Neighborhood

The threshold α tripartitions the neighborhood membership domain into certain positive, negative, and uncertain shadow regions, and thereby determines the structure of the shadowed neighborhoods. Improper thresholds will cause great membership loss and lead to over big or small uncertain regions of shadowed neighborhoods. A reasonable threshold should maintain the information of memberships when transforming neighborhood memberships into a shadowed neighborhood.

Suppose the membership function of a neighborhood is $\mu(x)$ and the neighborhood membership of any data instance $x_i \in X$ is $\mu(x_i) = \mu_i$, referring to the formula (8), the membership loss for transforming the neighborhood memberships into a shadowed set becomes

$$L(\alpha) = \lambda \cdot \left[\sum_{\mu_i \leq \alpha} \mu_i + \sum_{\mu_i \geq 1-\alpha} (1-\mu_i)\right] + \sum_{\alpha < \mu_i < 1-\alpha} |0.5 - \mu_i|. \tag{9}$$

Aiming to maintain the information in the transformation, the optimum threshold α^* should lead to the minimum membership loss,

$$\alpha^* = \arg\min_{\alpha} L(\alpha). \tag{10}$$

Based on the following piecewise representation of membership μ_i

$$u_i = \begin{cases} \mu_i, & \mu_i \leq 0.5 \\ 1-\mu_i, & \mu_i > 0.5 \end{cases} \tag{11}$$

we rewrite the neighborhood memberships of n data instances $\{\mu_1, \ldots, \mu_i, \ldots, \mu_n\}$ to $\{u_1, \ldots, u_i, \ldots, u_n\}$, $u_i \leq 0.5$ and reformulate the membership loss as

$$L(\alpha) = \lambda \cdot \sum_{u_i \leq \alpha} u_i + \sum_{u_i > \alpha} (0.5 - u_i). \tag{12}$$

$L(\alpha)$ consists of two parts, the first part denotes the membership loss in certain regions and the second part denotes the membership loss in uncertain region. Fixing the balance factor λ, the optimal threshold α^* of the minimum $L(\alpha)$ should tradeoff the two parts of membership loss.

Lemma 1: In the objective of membership loss $L(\alpha)$, for $\alpha \in [0, 0.5]$, $\lambda \cdot \sum_{u_i \leq \alpha} u_i$ is monotonically increasing and $\sum_{u_i > \alpha} (0.5 - u_i)$ is monotonically decreasing with respect to α. Therefore, the threshold α^* which leads to the minimum $L(\alpha)$ should tradeoff the membership loss of both certain and uncertain regions.

Based on the Lemma 1, we can infer the calculation of the optimal threshold to achieve the minimum membership loss $L(\alpha)$, see the following theorem.

Theorem 1: For a given $\lambda \in \mathbf{R}^+$, suppose $\alpha \in [0, 0.5]$, the membership loss $L(\alpha)$ achieves the minimum when $\alpha = \frac{0.5}{1+\lambda}$, i.e., the optimal threshold $\alpha^* = \arg\min_{\alpha} L(\alpha) = \frac{0.5}{1+\lambda}$.

Proof: $L(\alpha) = \lambda \cdot \sum_{u_i \leq \alpha} u_i + \sum_{u_i > \alpha} (0.5 - u_i)$, according to Lemma 1, in the objective of $L(\alpha)$, when α increases from 0 to 0.5, the membership loss of certain region $\lambda \cdot \sum_{u_i \leq \alpha} u_i$ monotonically increases and the increments grow as α increasing, in the meantime, the membership loss of uncertain region $\sum_{u_i > \alpha} (0.5 - u_i)$ monotonically decreases and the decrements gradually reduce. Therefore, the optimal threshold α^* leading to the minimum $L(\alpha^*)$ should tradeoff the growing loss increment of the certain region and the reducing loss decrement of the uncertain region.

Suppose $\alpha \in [0, 0.5]$ and ε is a small positive number. If there exists no membership value in the interval $(\alpha, \alpha + \varepsilon]$, we directly have $L(\alpha) = L(\alpha + \varepsilon)$, otherwise $\exists u_k, \alpha < u_k \leq \alpha + \varepsilon$. We use $diffL(\alpha)$ to denote the difference between the membership loss $L(\alpha)$ and $L(\alpha + \varepsilon)$ which can be also considered as the gradient of $L(\alpha)$ at α.

$$\begin{aligned} diffL(\alpha) &= L(\alpha + \varepsilon) - L(\alpha) \\ &= \lambda \cdot \sum_{u_i \leq \alpha + \varepsilon} u_i + \sum_{u_i > \alpha} (0.5 - u_i) \\ &\quad - \left[\lambda \cdot \sum_{u_i \leq \alpha} u_i + \sum_{u_i > \alpha + \varepsilon} (0.5 - u_i)\right] \end{aligned}$$

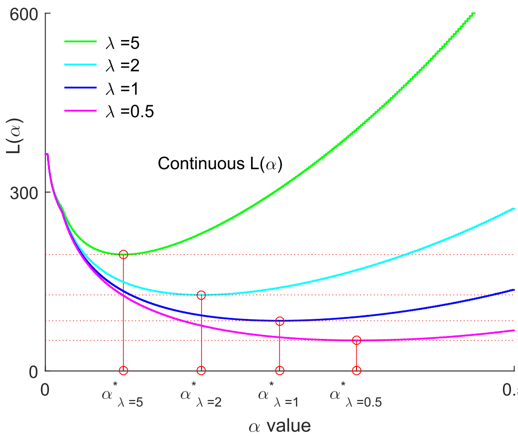

Fig. 4. Thresholding of the minimum $L(\alpha)$ on continuous membership.

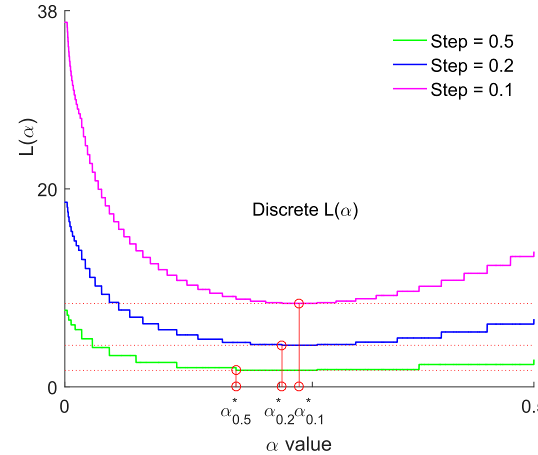

Fig. 5. Thresholding of the minimum $L(\alpha)$ on discrete memberships.

$$
\begin{aligned}
&= \lambda \cdot \left[\sum_{u_i \le \alpha+\varepsilon} u_i - \sum_{u_i \le \alpha} u_i\right] \\
&\quad + \left[\sum_{u_i > \alpha+\varepsilon} (0.5 - u_i) - \sum_{u_i > \alpha} (0.5 - u_i)\right] \\
&= \lambda \cdot \left[\sum_{u_i \le \alpha} u_i + u_k - \sum_{u_i \le \alpha} u_i\right] \\
&\quad + \left[\sum_{u_i > \alpha+\varepsilon} (0.5 - u_i) - \left(\sum_{u_i > \alpha+\varepsilon} (0.5 - u_i) + (0.5 - u_k)\right)\right] \\
&= \lambda \cdot u_k - (0.5 - u_k) \\
&= (1+\lambda) \cdot u_k - 0.5.
\end{aligned}
$$

From the formulas above, we know that the gradient $diffL(\alpha)$ is the sum of the membership loss variation in the certain and uncertain regions. Let $diffL(\alpha) = L(\alpha) - L(\alpha + \varepsilon) \le 0, diffL(\alpha) = (1+\lambda) \cdot u_k - 0.5 \le 0 \Rightarrow u_k \le \frac{0.5}{1+\lambda}$. Because $\alpha < u_k \le \alpha + \varepsilon, \alpha < u_k \le \frac{0.5}{1+\lambda}$ and thus $\forall \alpha \in [0, \frac{0.5}{1+\lambda})$, $diffL(\alpha) \le 0$. Similarly, $diffL(\alpha) \ge 0 \Rightarrow u_k \ge \frac{0.5}{1+\lambda}$, we have $\alpha + \varepsilon \ge u_k \ge \frac{0.5}{1+\lambda}$ and infer that $\forall \alpha \in [\frac{0.5}{1+\lambda}, 0.5]$, $diffL(\alpha) \ge 0$. Therefore, $L(\alpha)$ is monotonically decreasing in the interval $[0, \frac{0.5}{1+\lambda})$ and increasing in $[\frac{0.5}{1+\lambda}, 0.5]$ with respect to α. The gradient $diffL(\alpha) = 0 \Rightarrow \alpha^* = \frac{0.5}{1+\lambda}$, which is the optimum threshold to tradeoff the membership loss of the certain and uncertain regions and $L(\alpha^*)$ achieves the minimum membership loss. ■

According to Theorem 1, for a continuous neighborhood membership function, we set the optimal threshold $\alpha^* = \frac{0.5}{1+\lambda}$, and for the discrete neighborhood memberships, we adopt the closest membership value to $\frac{0.5}{1+\lambda}$ as the optimal threshold to construct shadowed neighborhoods. Fig. 4 presents the optimal thresholds for the continuous neighborhood membership function of Definition 2 under multiple λ values. Discretizing the continuous neighborhood membership function with multiple step lengths of 0.1, 0.2, and 0.5, we calculate the optimal thresholds for the three sets of discrete membership values and present the results in Fig. 5. It can be found that the optimal thresholds

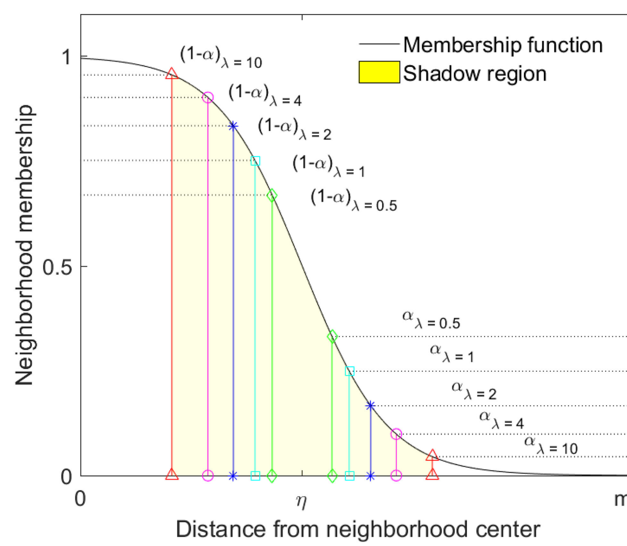

Fig. 6. Variation of neighborhood shadow against the balance factor λ.

obtained by Theorem 1 are effective to achieve the minimum membership loss for both continuous and discrete memberships.

Besides the optimal membership threshold, we further investigate the correlation between neighborhood shadows and the balance factor λ and infer the theorem as follows.

Theorem 2: The neighborhood shadow (uncertain boundary region) is monotonically increasing with respect to the balance factor λ of membership loss.

Proof: The neighborhood shadow is determined by the optimal threshold α^* and the size of shadow is denoted by the interval $(\alpha^*, 1-\alpha^*)$. $\forall \lambda_1, \lambda_2 \in \mathbf{R}^+, \lambda_1 \le \lambda_2, \alpha_1^* = \frac{0.5}{1+\lambda_1}$, $\alpha_2^* = \frac{0.5}{1+\lambda_2}$, thus we have $\alpha_2^* \le \alpha_1^*$ and infer that α^* monotonically decreases as λ increasing. Moreover, because $\lambda_1 \le \lambda_2 \Rightarrow \alpha_2^* \le \alpha_1^* \Rightarrow 1 - \alpha_2^* \ge 1 - \alpha_1^*$, the corresponding intervals satisfy $(\alpha_1^*, 1-\alpha_1^*) \subseteq (\alpha_2^*, 1-\alpha_2^*)$, which prove that the shadow size is monotonically increasing with respect to λ. ■

Fig. 6 illustrates the variation of shadow against the balance factor $\lambda = \{0.5, 1, 2, 4, 10\}$ and the shadow area gradually increases as λ increasing. As seen from the formula (12), the factor factor λ is used to tradeoff the membership losses of certain and uncertain regions in the transformation of shadowed neighborhood, and also can be viewed as the cost for changing the memberships in certain regions. Large values of λ indicate the great costs for reducing low memberships to certain 0 or elevating high memberships to certain 1. Therefore, the shadow area will be increased to include more instances as uncertain cases to reduce the costs of certain judgements. For the three-way classification with shadowed neighborhoods, we can control the rates of uncertain instances through adjusting the factor λ.

IV. Three-Way Classification With Shadowed Neighborhoods

Constructing a set of shadowed neighborhoods on labeled training data, we can implement a three-way classification method to classify unknown data instances into certain classes and uncertain case. The union of the shadowed neighborhoods of a class forms a tripartitioned approximation of the data distribution of the class. The classification of data instances is determined by the belongingness of the instances to the shadowed neighborhoods of different classes. To classify an instance x, we should first determine the regions of x in shadowed neighborhoods through thresholding its neighborhood memberships.

As shown in Theorem 1, the optimum threshold α^* of shadowed neighborhoods is determined by the factor λ which balance the costs of the membership losses on certain and uncertain regions. Therefore, we compute $\alpha^* = \frac{0.5}{1+\lambda}$ to threshold neighborhood memberships and partition the shadowed neighborhoods. Referring to Theorem 2, through setting λ, we can adjust the shadow regions of neighborhoods and the decision risk to suit the requirements of different classification tasks. For the cautious decision making, we can set high λ values to enlarge shadow regions of neighborhoods and thereby separate more uncertain instances for delayed decision making. For the efficient decision making which needs more automatic classifications, we can set low λ values to produce narrow shadow regions and lead to a few uncertain cases.

Suppose $\mu_k(x)$ is the membership of x to the kth neighborhood, POS_k, NEG_k and BND_k are the certain positive region, certain negative region, and the uncertain boundary region of the neighborhood, we distribute x into the three neighborhood regions in the following way.

$$\alpha < \mu_k(x) < 1-\alpha \Rightarrow x \in \mathrm{BND}_k$$
$$\mu_k(x) \leq \alpha \Rightarrow x \in \mathrm{NEG}_k$$
$$\mu_k(x) \geq 1-\alpha \Rightarrow x \in \mathrm{POS}_k.$$

With the high memberships $\geq 1-\alpha$, POS_k consists of the data instances certainly belonging to the kth neighborhood. NEG_k consists of the instances with the low memberships $\leq \alpha$, which are certainly beyond the neighborhood. BND_k consists of the uncertain instances locating in the neighborhood shadow area. Obtaining the neighborhood regions of x, we further define the following sets of neighborhood indexes to describe the region location of x to all the shadowed neighborhoods.

$$\begin{aligned}\mathrm{SNP}(x) &= \{k|x \in \mathrm{POS}_k\},\\ \mathrm{SNU}(x) &= \{k|x \in \mathrm{BND}_k\},\\ \mathrm{SNN}(x) &= \{k|x \in \mathrm{NEG}_k\}.\end{aligned}$$

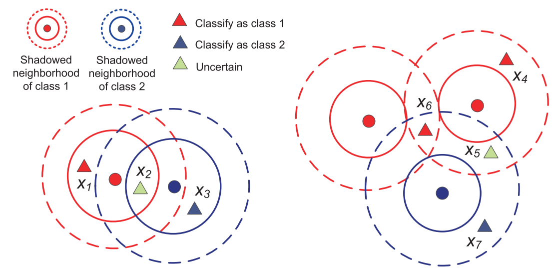

Fig. 7. 3-Way classification for instances within shadowed neighborhoods.

Obviously, $\mathrm{SNP}(x)$ is the set of the indexes of the neighborhoods whose positive regions containing the instance x, $\mathrm{SNU}(x)$ is the set of the indexes of the neighborhoods in which x locates in the uncertain boundary region, and $\mathrm{SNN}(x)$ denotes the set of neighborhoods excluding x. Given a set of neighborhoods $\mathbf{O} = \{O_1, \ldots O_k, \ldots O_K\}$, based on the region description of x provided by the neighborhood index sets, we can design a group of three-way classification rules to classify x in both conditions of x within and beyond the neighborhood set $\mathbf{O}$. In the classification rules, we adopt $class(O_k)$ to denote the class of the neighborhood O_k, i.e., the class of the neighborhood center x_k.

A. Classification Rules Within Shadowed Neighborhoods

For a data instance x locating within the neighborhoods of $\mathbf{O}$, we have $\exists O_k \in \mathbf{O}$, $\mu_k(x) > \alpha$, $|SNP(x)| \geq 1$ or $|SNU(x)| \geq 1$.

1) If $|\mathrm{SNP}(x)| = 1$, x certainly belongs to the class of the unique neighborhood in $\mathrm{SNP}(x)$.
2) If $|\mathrm{SNP}(x)| > 1$ and $\forall k_1, k_2 \in \mathrm{SNP}(x)$, $\mathrm{class}(O_{k_1}) = \mathrm{class}(O_{k_2})$, x certainly belongs to the class of the neighborhoods in $\mathrm{SNP}(x)$, otherwise if $\exists k_1, k_2 \in \mathrm{SNP}(x)$ and $\mathrm{class}(O_{k_1}) \neq \mathrm{class}(O_{k_2})$, x belongs to multiple neighborhoods of different classes with conflict and should be judged as an uncertain data instance.
3) If $|\mathrm{SNP}(x)| = 0$, $|\mathrm{SNU}(x)| > 0$, the major class of the neighborhoods in $\mathrm{SNU}(x)$ is C_m, $|\{k|k \in \mathrm{SNU}(x) \wedge \mathrm{class}(O_k) = C_m\}|/|\mathrm{SNU}(x)| \geq 60\%$, x belongs to the class C_m, otherwise if $|\{k|k \in \mathrm{SNU}(x) \wedge \mathrm{class}(O_k) = C_m\}|/|\mathrm{SNU}(x)| < 60\%$, x is judged as an uncertain data instance.

The within-neighborhood classification rules indicate that, if the shadowed neighborhoods whose positive regions containing x belong to the same class, we can certainly classify the instance, otherwise x belonging to heterogenous neighborhoods will lead to classification conflict and x should be considered as an uncertain instance. If x locates in the boundary regions (shadows) of multiple neighborhoods, we classify the instance through checking whether most of these neighborhoods belong to the same class. Fig. 7 illustrates the three-way classification rules for the instances within shadowed neighborhoods.

B. Classification Rules Beyond Shadowed Neighborhoods

For a data instance x beyond the neighborhood set $\mathbf{O}$, we have $\forall O_k \in \mathbf{O}$, $\mu_k(x) \leq \alpha$, $|SNP(x)| = 0$, $|SNU(x)| = 0$.

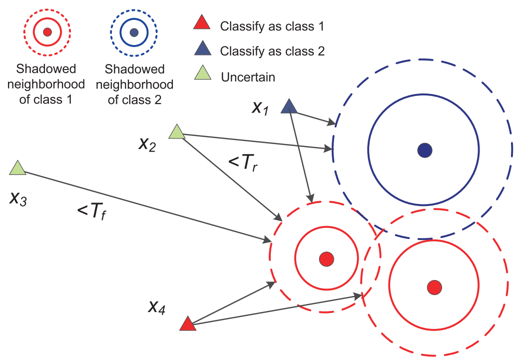

Fig. 8. 3-Way classification for instances beyond shadowed neighborhoods.

1) $\mu_f(x) = \max_{O_k \in \mathbf{O}}\{\mu_k(x)\}$, O_f is the nearest neighborhood of x, if $\mu_f(x) < T_f$, x is judged as an uncertain data instance.
2) $\mu_f(x) = \max_{O_k \in \mathbf{O}}\{\mu_k(x)\}$, $\mu_s(x) = \max_{O_k \in \mathbf{O}-\{O_f\}}\{\mu_k(x)\}$, O_f, O_s are the first and second nearest neighborhoods of x, if $\mu_f(x) \geq T_f$ and $class(O_f) = class(O_s)$, x belongs to the class of O_f and O_s.
3) $\mu_f(x) = \max_{O_k \in \mathbf{O}}\{\mu_k(x)\}$, $\mu_s(x) = \max_{O_k \in \mathbf{O}-\{O_f\}}\{\mu_k(x)\}$, O_f, O_s are the first and second nearest neighborhoods of x, if $\mu_f(x) \geq T_f$, $class(O_f) \neq class(O_s)$ and $1 - \mu_s(x)/\mu_f(x) \geq T_r$, x belongs to the class of O_f, otherwise if $1 - \mu_s(x)/\mu_f(x) < T_r$, x is judged as an uncertain data instance.

Different from the rules within neighborhoods, the three-way classification of the instances beyond neighborhoods depends on the distances between instances and neighborhoods. If the membership of x to its nearest neighborhood is too small and less than the threshold T_f, x is far from all the neighborhoods and should be considered as an uncertain instance. For the instances nearby neighborhoods, we determine the class of x according to its nearest two neighborhoods. If the two neighborhoods belong to the same class, we can perform the certain classification. Otherwise we further check the difference between the memberships of x to its first and second nearest neighborhoods of different classes. If the membership difference is less than the threshold T_r, which means the distances from x to the referenced two neighborhoods are similar, the class inconsistency of the two neighborhoods leads to the uncertain judgement of x. If the membership difference is big enough ($\geq T_r$), we can certainly determine the class of x referring to only the nearest neighborhood. In the experiments, we set $T_f = 0.05$ and $T_r = 0.1$ as default. The three-way classification rules beyond shadowed neighborhoods are illustrated in Fig. 8.

Summarizing the three-way classification rules within and beyond neighborhoods, we implement a three-way classification algorithm with shadowed neighborhoods (3WC-SNB). The detailed flow of the algorithm is presented in Algorithm 1.

Utilizing Algorithm 1 to classify a set of data instances X, the number of instances $|X| = n$, it is required to calculate the memberships of each instance to K neighborhoods. In the algorithm implementation, we build up a $n \times K$ matrix of instance-neighborhood memberships to achieve this. Thus the computational complexity of the test phase is $O(n \times K)$. Because $K \ll n$, the classification based on neighborhoods is more efficient than the neighbor-based classification. In the training phase, the construction of neighborhoods needs to search the nearest homogeneous and heterogeneous neighbors of each instance, thus the computational complexity of neighborhood construction is $O(n^2)$. We can further speed up the neighborhood construction under divide-and-conquer strategies, such as using K-Dimensional tree to speed up the neighbor searching. Moreover, extending neighborhoods to shadowed ones requires to compute the membership threshold for each neighborhood and thus needs $O(K)$ calculations. The computational complexity in training phase is summarized as $O(n^2 + K) \approx O(n^2)$.

V. Experimental Results

Different from the certain classification methods, the three-way classification method based on shadowed neighborhoods (3WC-SNB) classifies data instances into certain classes and the uncertain case, which is is helpful to avoid the farfetched classification of uncertain (or challenging) instances and thereby reduce the classification risk. To validate this, we implement three tests in the experiment. In the first test, we compare the certain classification with neighborhoods and the three-way classification with shadowed neighborhoods on noisy data to verify the effectiveness of 3WC-SNB for uncertain data classification. The second test validates the superiority of 3WC-SNB in low-risk classification through comparing the proposed method with other typical certain classification methods. In the third test, we further compare 3WC-SNB with the 3WD method based on attribute reduction [28] to validate the superiority of the proposed method for numeric data analysis. Focusing on the risk of classification, we collect 13 datasets in the areas of medicine and economics from the University of California Irvine machine learning data repository to implement the experiment. For all the tests in the experiment, tenfold cross validation is performed on each data set. The descriptions of the adopted data sets are given in Table I.

We set the minor class as the positive class for each data set. For an example, in the breast cancer datasets, the class of "malignant" will be set as the positive class. Suppose the number of the positive-class instances is P and the number of the negative-class instances is N, TP, and FP denote the numbers of true positive and false positive classified instances, TN and FN denote the numbers of true negative and false negative classified instances. To overall evaluate the classification methods, we adopt the measures of *accuracy*, *precision*, *recall rate*, *F1 score*, *ratio of uncertain instances* (*UR*), and *classification cost* as the evaluation criteria. The calculations of these measures are listed as follows.

$$\begin{aligned}
&\text{Accuracy} = (TP + TN)/(P + N),\\
&\text{Precision} = TP/(TP + FP),\\
&\text{Recall Rate} = TP/P,\\
&\text{F1Score} = 2 \cdot \text{Precsion} \cdot \text{Recall}/(\text{Precsion+Recall}),\\
&\text{UR} = |\{x | x \in X_{\text{test}} \wedge \text{class}(x) = \text{uncertain}\}|/|X_{\text{test}}|,\\
&\text{Cost} = C_{NP} \cdot \tfrac{FP}{P+N} + C_{PN} \cdot \tfrac{FN}{P+N} + C_U \cdot \text{UR}.
\end{aligned}$$

Algorithm 1: Three-Way Classification With Shadowed Neighborhoods (3WC-SNB).

Input: K shadowed neighborhoods with the optimized thresholds
$\mathbf{N} = \{N^{\alpha}_{\mu_1}, N^{\alpha}_{\mu_2}, \ldots, N^{\alpha}_{\mu_K}\}$;
Unknown data instance x;
Output: Three-way classification result of x, $class(x)$;
Initialize SNP(x), SNU$(x) \rightarrow \emptyset$;
Compute neighborhood memberships of x for K neighborhoods $\{\mu_1(x), \mu_2(x), \ldots, \mu_K(x)\}$ according to the formula (6);
//Determine the region of x according to neighborhood memberships
for each shadowed neighborhood $N^{\alpha}_{\mu_k}$ **do**
  if $\mu_k(x) \geq 1 - \alpha$ **then**
    $N^{\alpha}_{\mu_k}(x) = 1$, SNP$(x)$ = SNP$(x) \cup \{k\}$;
  else
    if $\mu_k(x) > \alpha$ **then**
      $N^{\alpha}_{\mu_k}(x) = 0.5$, SNU$(x)$ = SNU$(x) \cup \{k\}$;
    end if
  end if
end for
//Instance x in positive regions of neighborhoods
if $|$SNP$(x)| \geq 1$ **then**
  Obtain the classes C_{SNP} of the neighborhoods in SNP;
  if $|C_{\text{SNP}} : \{C_p\}| = 1$ **then**
    class$(x) = C_p$;
  else
    class(x) = uncertain;
  end if
else
  //Instance x in boundary regions of neighborhoods
  if $|$SNU$(x)| \geq 1$ **then**
    Obtain the major class C_m of the neighborhoods in $SNU(x)$;
    if $\frac{|\{k | k \in \text{SNU}(x) \wedge \text{class}(O_k) = C_m\}|}{|\text{SNU}(x)|} \geq 60\%$ **then**
      class$(x) = C_m$;
    else
      class(x) = uncertain;
    end if
  end if
end if
//Instance x beyond neighborhoods
if $|$SNP$(x)| = 0$ and $|$SNU$(x)| = 0$ **then**
  Compute the memberships of x for the first and second nearest neighborhoods O_f, O_s,
  $\mu_f(x) = \max_{1 \leq k \leq K}\{\mu_k(x)\}$,
  $\mu_s(x) = \max_{1 \leq k \leq K \wedge k \neq f}\{\mu_k(x)\}$;
  if $\mu_f(x) < T_f$ **then**
    class(x) = uncertain;
  else
    if class(O_f) = class(O_s) **then**
      class(x) = class(O_f);
    else
      if $1 - \mu_s(x)/\mu_f(x) \geq T_r$ **then**
        class(x) = class(O_f);
      else
        class(x) = uncertain;
      end if
    end if
  end if
end if
Return class(x).

TABLE I
EXPERIMENTAL DATASETS

Data sets	*Feature*	*Instance*	*Class Ratio*	*Type*
Appendicitis	7	106	20% vs. 80%	Numerical
Banknote Authentication	4	1372	44% vs. 56%	Numerical
Blood Transfusion Service Center	4	748	24% vs. 76%	Numerical
Wisconsin Original Breast Cancer	9	699	34% vs. 66%	Numerical
Fertility	9	100	12% vs. 88%	Numerical
German Credit	24	1000	30% vs. 70%	Numerical
Haberman's Survival	3	306	26% vs. 74%	Numerical
Indian Liver Patients	10	583	29% vs. 71%	Numerical
Mammographic Mass	5	961	46% vs. 54%	Numerical
Thoracic Surgery	16	470	15% vs. 85%	Numerical
Wisconsin Diagnostic Breast Cancer	30	569	37% vs. 63%	Numerical
Wisconsin Prognostic Breast Cancer	33	198	24% vs. 76%	Numerical
Sensorless Drive Diagnosis	49	58509	c5 vs. c6	Numerical

In the cost measure, the cost of correct classification is zero, C_{NP}, C_{PN}, C_U denote the costs of false-positive classification, false-negative classification, and the classification of uncertain instances, respectively. For the medical and economic data, misclassifying positive instances (of minor class) as negative ones causes more costs than the misclassification of negative instances, such as classifying malignant tumors as benign will suffer more risk than judging benign tumors as malignant. The classification of uncertain instances will delay the decision making and thus has less cost than false-positive and false-negative classifications. Therefore, we set $C_{PN}/C_{NP}/C_U = 5/1/0.5$ in the following tests.

A. Test of Uncertain Data Classification

To validate the effectiveness of the proposed shadowed neighborhoods for uncertain data classification, we expect to apply 3WC-SNB method to classify the data with multilevel uncertainty. The inconsistency between training data and test data gives rise to the uncertainty in classification process, thus we produce the uncertain instances for classification through adding multilevel noise to test data. Specifically, we randomly change the class labels of partial instances from 0% to 50% in the test dataset and produce the test dataset with multilevel label noise.

We construct both the shadowed neighborhoods and traditional neighborhoods [68] on the same training datasets and perform the 3WC-SNB and the certain classification based on the nearest neighborhoods (2WC-NB) on the test datasets with multilevel noise. Fig. 9 shows the TP rate (TP/P) and FN rate (FN/P) of 3WC-SNB and 2WC-NB against the noise level from 0% to 50%. We can find that on different noise levels, 3WC-SNB

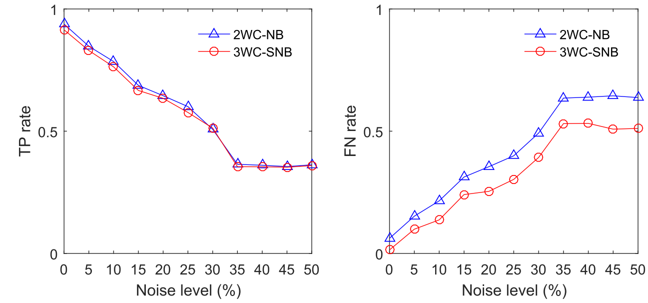

Fig. 9. TP rates and FN rates of 3WC-SNB and 2WC-NB on noisy data.

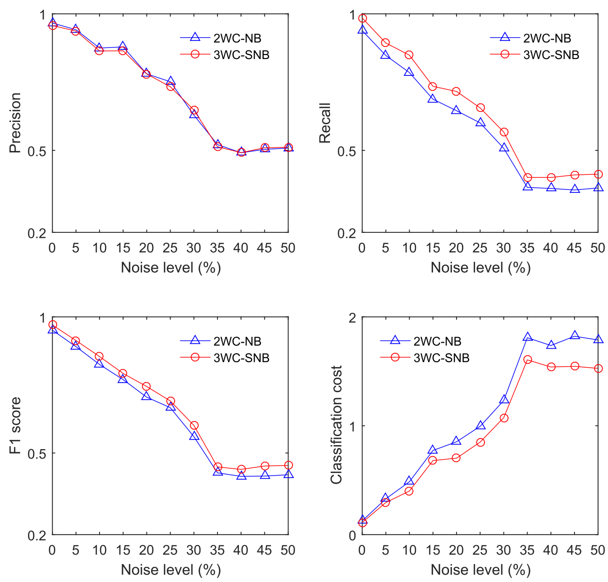

Fig. 10. Classification results of 3WC-SNB and 2WC-NB on noisy data.

and 2WC-NB achieve the very similar TP rates but 3WC-SNB generates less FN rates than 2WC-NB. This indicates that 3WC-SNB can correctly classify the typical positive instances as 2WC-NB, and in the meantime, reduce the misclassifications of positive instances to negative class through separating uncertain instances. We can also find that the gap of FN rate between 3WC-SNB and 2WC-NB is widening as the noise level increasing, which means the 3WC-SNB tends to recognize more uncertain instances when the test datasets contain more noise.

Due to the similar TP rates and less FN rates, the 3WC-SNB achieves more precise classification results than the certain classification based on traditional neighborhoods. Fig. 10 illustrates the precision, recall rates, classification costs, and F1 scores of the classification results produced by 3WC-SNB and 2WC-NB on the multilevel noisy datasets. More detailed evaluations of the classification results are presented in Table II. Comparing with 2WC-NB, 3WC-SNB produces the higher recall rates and F1 scores, and the lower classification costs. Especially for the datasets with heavy noise (much uncertainty), the proposed three-way classification method can avoid the serious misclassifications and greatly reduces the classification risk.

TABLE II
CLASSIFICATION RESULTS ON MULTILEVEL NOISY DATA

Noise	*Methods*	*TP*	*FN*	*Cost*	*Acc*	*Prec*	*Rec*	*F1*
0%	3WC-SNB	0.91	0.01	0.11	0.84	0.96	0.98	0.97
	2WC-NB	0.94	0.06	0.13	0.96	0.97	0.94	0.95
5%	3WC-SNB	0.83	0.10	0.29	0.80	0.94	0.89	0.91
	2WC-NB	0.85	0.15	0.33	0.92	0.94	0.85	0.89
10%	3WC-SNB	0.76	0.14	0.40	0.76	0.86	0.85	0.85
	2WC-NB	0.78	0.22	0.49	0.87	0.87	0.78	0.83
15%	3WC-SNB	0.67	0.24	0.68	0.70	0.86	0.73	0.79
	2WC-NB	0.69	0.31	0.77	0.81	0.88	0.69	0.77
20%	3WC-SNB	0.63	0.25	0.70	0.67	0.78	0.72	0.74
	2WC-NB	0.65	0.35	0.85	0.77	0.78	0.65	0.71
25%	3WC-SNB	0.58	0.30	0.85	0.63	0.73	0.66	0.69
	2WC-NB	0.60	0.40	1.00	0.73	0.75	0.60	0.67
30%	3WC-SNB	0.51	0.39	1.07	0.56	0.65	0.57	0.60
	2WC-NB	0.51	0.49	1.23	0.65	0.63	0.51	0.56
35%	3WC-SNB	0.35	0.53	1.61	0.41	0.51	0.40	0.45
	2WC-NB	0.36	0.64	1.81	0.50	0.52	0.36	0.43
40%	3WC-SNB	0.35	0.53	1.54	0.42	0.49	0.40	0.44
	2WC-NB	0.36	0.64	1.74	0.51	0.49	0.36	0.41
45%	3WC-SNB	0.35	0.51	1.55	0.43	0.51	0.41	0.45
	2WC-NB	0.36	0.64	1.82	0.49	0.50	0.36	0.42
50%	3WC-SNB	0.36	0.51	1.53	0.43	0.51	0.41	0.45
	2WC-NB	0.36	0.64	1.79	0.50	0.51	0.36	0.42

TABLE III
CLASSIFICATION RESULTS OF OF DIFFERENT CLASSIFICATION METHODS

Methods	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	57.81	81.17	82.85	81.17	81.07
Decision-Tree (J48)	42.46	83.89	82.07	83.75	82.60
SVM	45.87	82.75	85.64	85.64	85.64
Cost-sensitive Bayes	47.44	79.38	82.48	79.29	78.67
Cost-sensitive J48	24.20	81.25	81.91	81.25	85.37
Cost-sensitive Bayes Net	27.71	79.95	81.23	79.97	83.39
3WC-SNB	**20.52**	**81.20**	**87.95**	**92.44**	**89.26**

The bold entities indicate the experimental results generated by the method proposed in this article.

B. Comparison With Certain Classification Methods

The second test overall evaluates the proposed shadowed-neighborhood-based uncertain classification method through comparing with multiple kinds of certain classification methods. We compare 3WC-SNB method with three elegant classification methods, namely, naive Bayes, support vector machine (SVM), and decision trees (J48) [70]. Moreover, focusing on the evaluation of classification risk, we also compare the proposed method with other three typical cost-sensitive classification methods, namely, cost-sensitive Bayes, cost-sensitive decision trees, and cost-sensitive Bayes net [71]. Fig. 11 and Table III present the average classification results on all the test datasets for each classification method and the details are listed in the appendix.

From the experimental results, we find that comparing with the certain classification methods, the proposed uncertain method generally produces lower classification accuracy. This is because that the uncertain data instances without class labels should not be counted in the calculation of accuracy. However, in contrast to all the certain classification methods, 3WC-SNB achieves higher recall rates and F1 scores, and thereby induces the lower classification costs. Only considering the classification error, SVM and decision trees produce precise classification results but suffer too much classification costs. Involving risks of misclassifications in classification process, the cost-sensitive methods reduce the classification costs but over classify data instances into the more risky class. Different from the cost-sensitive methods forcing to classify instances into the classes of high risks, 3WC-SNB reduces classification costs through

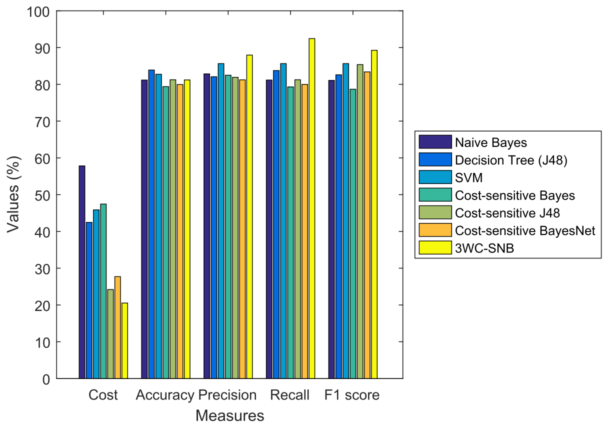

Fig. 11. Comparison of classifications of different methods.

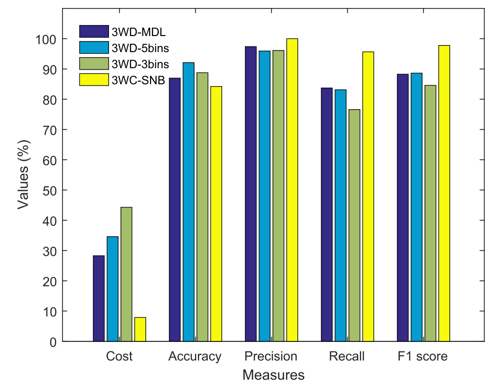

Fig. 12. Comparison of classifications of 3WC-SNB and discretized 3WD.

TABLE IV
CLASSIFICATION RESULTS OF 3WC-SNB AND 3WD WITH DISCRETIZATION

Methods	*TP* (%)	*FN* (%)	*UR* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
3WD-MDL	76.36	13.39	7.42	28.28	86.97	97.39	83.71	88.27
3WD-5bins	83.13	16.87	0	34.58	92.1	95.93	83.13	88.62
3WD-3bins	74.6	22.28	1.4	44.29	88.75	96.11	76.59	84.58
3WC-SNB	**95.65**	**0**	**15.79**	**7.89**	**84.21**	**100**	**95.65**	**97.78**

The bold entities indicate the experimental results generated by the method proposed.

delaying the challenging classifications of a limited number of uncertain instances. In general, the uncertain classification method based on shadowed neighborhoods outperforms the certain classification methods and is effective to reduce the classification costs.

C. Comparison With Three-Way Decision Method

Besides the certain classification methods, we also compare the proposed three-way classification method 3WC-SNB with another elegant 3WD method which is constructed based on probabilistic attribute reduction [28]. Probabilistic attribute reduction formulates three-way decision rules through constructing the probabilistic attribute reducts, which partition data instances into positive, negative, and boundary regions for a given class. Different from the shadowed neighborhoods constructed on the numerical data (or mixed-type data), probabilistic attribute reduction is used to extract decision rules from symbolic datasets and requires data discretization for numerical data analysis. Moreover, different from 3WC-SNB estimates the membership threshold α^* through optimizing the neighborhood shadow, 3WD method utilizes a pair of parameters $(\alpha, \beta) \in [0, 1], \alpha < \beta$ to threshold the memberships and thereby tripartitions data instances into certain classes and uncertain case.

Performing 3WD method to classify the numerical data, we apply both the supervised multiinterval discretization method (MDL) and the unsupervised Equal-width Discretization method (5 bins and 3 bins) [72] to discretize the numerical attribute values of the test datasets, and set the threshold parameters $\alpha = 0.5$, $\beta = 0.8$ as default. Fig. 12 illustrates the classification results of 3WC-SNB and 3WD with different discretization strategies and Table IV presents the details. The experimental results indicate that the classification based on 3WD is not stable for different discretization methods. The preprocessing of discretization may bring about the information loss and thus make the three-way decision rules produce imprecise classification results. Besides the effects of data discretization, the classification of 3WD is also sensitive to the threshold parameter setting. The quality of the decision rules generated by the attribute reducts relies on the predefined α, β adopted in the probabilistic attribute reduction. Depending on the superiorities of shadowed neighborhoods in numerical data processing, and the optimization of thresholding parameter, the proposed 3WC-SNB method achieves stable and precise classification results.

VI. CONCLUSION

In this article, we propose a novel shadowed set to construct shadowed neighborhoods for uncertain data classification. Specifically, the proposed shadowed sets utilize a step function to map neighborhood memberships to the set of typical certain and uncertain membership values and thereby partition a neighborhood into the certain positive, negative, and uncertain boundary regions. The threshold parameter in the step function for constructing shadowed neighborhoods was optimized through minimizing the membership loss in the shadowed mapping. Based on the constructed shadowed neighborhoods, we also design three-way classification rules and thereby implement a three-way classification algorithm to distinguish data instances into certain classes and uncertain case. Experiments verify the superiorities of the proposed three-way method for classifying uncertain data and reducing classification risks.

Our future works may include the following issues. First, the memberships of shadowed neighborhood are computed based on distances, and thereby model the ball-shaped data distribution well but are not flexible enough for complex data distributions. To handle the diverse data, we should consider the distributions in local regions to compute neighborhood memberships. Second, we will further investigate the optimization strategy of shadowed neighborhoods through involving the classification error (or costs) in the objective. The final issue is that, we adopt Euclidean distances to construct the neighborhoods and compute the memberships, but this distance metric will not be effective for high-dimensional data. Therefore the feature reduction and kernel methods will be further involved in the construction of shadowed neighborhoods.

APPENDIX

TABLE V
CLASSIFICATION RESULTS ON DATASET "APPENDICITIS"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	85.80	44.34	85.85	86.10	85.80	86.00
Decision-Tree (J48)	85.80	33.02	85.85	84.90	85.80	85.10
SVM	75.01	50.07	83.33	75.01	75.01	75.01
Cost-sensitive Bayes	89.60	25.47	89.62	89.20	89.60	89.30
Cost-sensitive J48	79.20	32.08	79.24	73.00	79.20	73.80
Cost-sensitive Bayes Net	80.20	19.81	80.19	80.20	80.20	89.00
3WC-SNB	**100.00**	**5.00**	**90.00**	**100.00**	**100.00**	**100.00**

The bold entities indicate the experimental results generated by the method proposed.

TABLE VI
CLASSIFICATION RESULTS ON DATA SET "BANKNOTE"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	84.30	41.98	84.25	84.30	84.30	84.20
Decision-Tree (J48)	98.50	4.66	98.54	98.50	98.50	98.50
SVM	100.00	0.00	100.00	100.00	100.00	100.00
Cost-sensitive Bayes	79.80	28.06	79.81	82.80	79.80	78.80
Cost-sensitive J48	98.01	3.72	98.03	98.10	98.00	98.00
Cost-sensitive Bayes Net	83.70	21.57	83.67	86.10	83.70	83.10
3WC-SNB	**93.30**	**12.77**	**83.21**	**85.37**	**93.33**	**89.17**

The bold entities indicate the experimental results generated by the method proposed.

TABLE VII
CLASSIFICATION RESULTS ON DATASET "BLOOD"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	75.40	47.06	75.40	71.00	75.40	71.60
Decision-Tree (J48)	77.80	56.95	77.81	76.40	77.80	76.90
SVM	78.67	95.45	64.84	78.67	78.67	78.67
Cost-sensitive Bayes	76.70	32.89	76.74	72.50	76.70	70.60
Cost-sensitive J48	76.20	23.81	76.20	76.20	76.20	86.50
Cost-sensitive Bayes Net	74.46	25.58	77.88	75.70	74.46	85.54
3WC-SNB	**76.19**	**23.33**	**65.33**	**84.21**	**76.19**	**80.77**

The bold entities indicate the experimental results generated by the method proposed.

TABLE VIII
CLASSIFICATION RESULTS ON DATASET "WOBC"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	96.00	16.66	95.99	96.20	96.00	96.00
Decision-Tree (J48)	94.60	16.88	94.56	94.60	94.60	94.60
SVM	97.83	9.09	97.10	97.83	97.83	97.83
Cost-sensitive Bayes	95.70	16.88	95.71	95.80	95.70	95.70
Cost-sensitive J48	92.00	15.45	91.99	92.10	92.00	91.80
Cost-sensitive Bayes Net	96.90	10.59	96.85	96.90	96.90	96.90
3WC-SNB	**97.73**	**7.35**	**98.53**	**100.00**	**97.73**	**98.85**

The bold entities indicate the experimental results generated by the method proposed.

TABLE IX
CLASSIFICATION RESULTS ON DATASET "FERTILITY"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	88.05	60.00	88.05	88.05	88.05	93.60
Decision-Tree (J48)	85.00	63.00	85.00	77.10	85.00	80.90
SVM	85.71	75.00	75.00	85.71	85.71	85.71
Cost-sensitive Bayes	75.00	53.00	75.00	82.60	75.00	78.10
Cost-sensitive J48	78.00	54.00	78.00	82.10	78.00	79.80
Cost-sensitive Bayes Net	53.00	59.00	53.00	84.40	53.00	60.70
3WC-SNB	**100.00**	**10.05**	**90.08**	**90.08**	**100.00**	**94.74**

The bold entities indicate the experimental results generated by the method proposed.

TABLE X
CLASSIFICATION RESULTS ON DATASET "GERMANCREDIT"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	75.70	61.90	75.70	74.70	75.70	74.90
Decision-Tree (J48)	73.90	68.90	73.90	72.90	73.90	73.20
SVM	83.64	20.69	79.35	83.64	83.64	83.64
Cost-sensitive Bayes	73.40	39.00	73.40	72.10	73.40	68.20
Cost-sensitive J48	70.10	30.05	71.00	70.10	70.10	82.40
Cost-sensitive Bayes Net	71.18	29.95	70.33	70.33	71.18	81.10
3WC-SNB	**81.69**	**30.50**	**61.89**	**77.33**	**81.69**	**79.45**

The bold entities indicate the experimental results generated by the method proposed.

TABLE XI
CLASSIFICATION RESULTS ON DATASET "HABERMAN"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	74.80	42.16	74.84	71.50	74.80	70.30
Decision-Tree (J48)	71.90	66.01	71.89	69.00	71.90	69.80
SVM	82.14	109.09	69.70	82.14	82.14	82.14
Cost-sensitive Bayes	74.20	37.58	74.18	70.10	74.20	67.90
Cost-sensitive J48	73.50	26.47	73.52	73.50	73.50	84.70
Cost-sensitive Bayes Net	75.33	30.56	71.32	73.50	75.33	81.40
3WC-SNB	**100.00**	**19.35**	**80.65**	**80.65**	**100.00**	**89.29**

The bold entities indicate the experimental results generated by the method proposed.

TABLE XII
CLASSIFICATION RESULTS ON DATASET "INDIAN LIVER PATIENTS"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	55.70	215.78	55.75	79.20	55.70	56.00
Decision-Tree (J48)	69.10	79.76	68.95	66.90	69.10	67.60
SVM	51.35	147.95	50.68	51.35	51.35	51.35
Cost-sensitive Bayes	56.90	208.40	56.96	78.80	56.90	57.50
Cost-sensitive J48	71.40	28.64	71.36	71.40	71.40	83.30
Cost-sensitive Bayes Net	70.50	35.68	70.49	59.80	70.50	60.20
3WC-SNB	**97.56**	**37.93**	**68.97**	**70.18**	**97.56**	**81.63**

The bold entities indicate the experimental results generated by the method proposed.

TABLE XIII
CLASSIFICATION RESULTS ON DATASET "MOGRAPHIC"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	82.50	61.60	82.52	82.90	82.50	82.50
Decision-Tree (J48)	82.40	51.30	82.41	82.40	82.40	82.40
SVM	95.45	10.53	96.77	95.45	95.45	95.45
Cost-sensitive Bayes	82.00	44.22	81.99	82.20	82.00	81.90
Cost-sensitive J48	76.00	34.03	75.96	79.90	76.00	74.70
Cost-sensitive Bayes Net	82.20	34.44	82.20	83.20	82.20	81.90
3WC-SNB	**63.41**	**45.11**	**77.24**	**76.47**	**63.41**	**79.93**

The bold entities indicate the experimental results generated by the method proposed.

TABLE XIV
CLASSIFICATION RESULTS ON DATASET "THORACIC"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	77.90	65.53	77.87	78.10	77.90	78.00
Decision-Tree (J48)	82.80	73.40	82.77	75.80	82.80	78.30
SVM	89.66	48.65	83.78	89.66	89.66	89.66
Cost-sensitive Bayes	63.80	57.45	63.83	81.00	63.80	69.00
Cost-sensitive J48	72.80	62.13	72.77	79.00	72.80	75.30
Cost-sensitive Bayes Net	83.00	74.05	82.98	75.30	83.00	78.20
3WC-SNB	**97.50**	**15.96**	**82.98**	**84.78**	**97.50**	**90.70**

The bold entities indicate the experimental results generated by the method proposed.

TABLE XV
CLASSIFICATION RESULTS ON DATASET "WDBC"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	93.00	22.50	92.97	93.00	93.00	93.00
Decision-Tree (J48)	93.30	17.22	93.32	93.40	93.30	93.30
SVM	88.89	20.69	93.10	88.89	88.89	88.89
Cost-sensitive Bayes	93.00	21.79	92.97	93.00	93.00	93.00
Cost-sensitive J48	94.00	12.30	94.02	94.30	94.00	94.10
Cost-sensitive Bayes Net	94.90	12.83	94.90	95.00	94.90	94.90
3WC-SNB	**100.00**	**9.65**	**85.96**	**100.00**	**100.00**	**93.33**

TABLE XVI
CLASSIFICATION RESULTS ON DATASET "WPBC"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	67.20	119.70	67.17	72.10	67.20	68.90
Decision-Tree (J48)	75.80	68.69	75.76	75.10	75.80	75.40
SVM	85.71	60.01	80.01	85.71	85.71	85.71
Cost-sensitive Bayes	72.70	79.80	72.73	72.30	72.70	72.50
Cost-sensitive J48	71.31	25.74	76.26	76.30	71.31	86.75
Cost-sensitive Bayes Net	76.40	36.74	76.26	76.30	76.40	86.50
3WC-SNB	**94.44**	**32.50**	**85.00**	**94.44**	**94.44**	**94.44**

The bold entities indicate the experimental results generated by the method proposed.

TABLE XVII
CLASSIFICATION RESULTS ON DATASET "SDD"

Methods	*TP* (%)	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
Naive Bayes	98.90	0.37	98.90	99.90	98.90	98.90
Decision-Tree (J48)	97.93	0.20	99.90	99.90	97.93	97.90
SVM	99.27	1.99	99.34	99.27	99.27	99.27
Cost-sensitive Bayes	97.97	0.19	98.89	99.90	97.97	99.90
Cost-sensitive J48	98.91	0.17	98.93	98.87	98.91	98.91
Cost-sensitive Bayes Net	99.19	0.23	98.97	99.17	99.19	99.17
3WC-SNB	**99.82**	**0.15**	**99.00**	**100.00**	**99.82**	**99.41**

The bold entities indicate the experimental results generated by the method proposed.

REFERENCES

[1] R. D. Short and K. Fukunaga, "Optimal distance measure for nearest neighbor classification," *IEEE Trans. Inf. Theory*, vol. 27, no. 5, pp. 622–627, Sep. 1981.

[2] W. P. Ding, C. T Lin, and Z. H. Cao, "Shared nearest-neighbor quantum game-based attribute reduction with hierarchical coevolutionary spark and its application in consistent segmentation of neonatal cerebral cortical surfaces," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 7, pp. 2013–2027, Jul. 2019.

[3] W. P. Ding, C. T. Lin, and Z. H. Cao, "Deep neuro-cognitive coevolution for fuzzy attribute reduction by quantum leaping PSO with nearest-neighbor memeplexes," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2744–2757, Jul. 2019.

[4] M. Muja and D. G. Lowe, "Scalable nearest neighbor algorithms for high dimensional data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 11, pp. 2227–2240, Nov. 2014.

[5] A. Owen, "A neighborhood-based classifier for LANDSAT data," *Can. J. Statist.*, vol. 12, pp. 191–200, 1984.

[6] H. Wang, "Nearest neighbors by neighborhood counting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 6, pp. 942–953, Jun. 2006.

[7] Q. H. Hu, W. Pedrycz, D. Yu, and J. Lang, "Selecting discrete and continuous features based on neighborhood decision error minimization," *IEEE Trans. Syst., Man, Cybern.-Part B*, vol. 40, no. 1, pp. 137–150, Feb. 2010.

[8] C. Z. Wang, Q. H. Hu, X. Z. Wang, D. Chen, Y. H. Qian, and Z. Dong, "Feature selection based on neighborhood discrimination index," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 7, pp. 2986–2999, Jul. 2018.

[9] Q. H. Hu, D. R. Yu, and Z. X. Xie, "Neighborhood classifier," *Expert Syst. Appl.*, vol. 34, pp. 866–876, 2008.

[10] X.D. Yue, Y. F. Chen, D. Q. Miao, and J. Qian, "Tri-partition neighborhood covering reduction for robust classification," *Int. J. Approx. Reasoning*, vol. 83, pp. 371–384, 2017.

[11] D. Wettschereck and T. G. Dieterich, "An experimental comparison of the nearest neighbor and nearest-hyper-rectangle algorithms," *Mach. Learn.*, vol. 19, pp. 5–27, 1995.

[12] H. Fujita, T. R. Li, and Y. Y. Yao, "Advances in three-way decisions and granular computing," *Knowl.-Based Syst.*, vol. 91, pp. 1–3, 2016.

[13] Y. F. Chen, X. D. Yue, H. Fujita, and S. Y. Fu, "Three-way decision support for diagnosis on focal liver lesions," *Knowl.-Based Syst.*, vol. 127, pp. 85–99, 2017.

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

[15] W. Pedrycz and G. Vukovich, "Granular computing with shadowed sets," *Int. J. Intell. Syst.*, vol. 17, no. 2, pp. 173–197, 2002.

[16] A. Bargiela and W. Pedrycz, "Toward a theory of granular computing for human-centered information processing," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 2, pp. 320–330, Apr. 2008.

[17] Z. Pawlak, "Rough sets," *Int. J. Inf. Comput. Sci.*, vol. 11, no. 5, pp. 314–356, 1982.

[18] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning About Data*. Boston, MA, USA: Kluwer, 1991.

[19] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, 1965.

[20] L. A. Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic," *Fuzzy Sets Syst.*, vol. 90, no. 2, pp. 111–127, 1997.

[21] D. Dubois and H. Prade, "Rough fuzzy sets and fuzzy rough sets," *Int. J. General Syst.*, vol. 17, pp. 191–209, 1990.

[22] S. Nanda and S. Majumdar, "Fuzzy rough sets," *Fuzzy Sets Syst.*, vol. 45, no. 2, pp. 157–160, 1992.

[23] R. Jensen and Q. Shen, "Fuzzy-rough sets assisted attribute selection," *IEEE Trans. Fuzzy Syst.*, vol. 15, no. 1, pp. 73–89, Feb. 2007.

[24] E. Ramentol *et al.*, "IFROWANN: Imbalanced fuzzy-rough ordered weighted average nearest neighbor classification," *IEEE Trans. Fuzzy Syst.*, vol. 23, no. 5, pp. 1622–1637, Oct. 2015.

[25] W. P. Ding, C. T. Lin, M. Prasad, Z. Cao, and J. Wang, "A layered-coevolution-based attribute-boosted reduction using adaptive quantum behavior PSO and its consistent segmentation for neonates brain tissue," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 3, pp. 1177–1191, Jun. 2018.

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

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

[28] Y. Y. Yao, "Three-way decisions with probabilistic rough sets," *Inf. Sci.*, vol. 180, pp. 341–353, 2010.

[29] Y. Y. Yao, "The superiority of three-way decisions in probabilistic rough set models," *Inf. Sci.*, vol. 181, pp. 1080–1096, 2011.

[30] A. M. Radzikowska and E. E. Kerre, "A comparative study of fuzzy rough sets," *Fuzzy Sets Syst.*, vol. 126, no. 2, pp. 137–155, 2002.

[31] W. Z. Wu, J. S. Mi, and W. X. Zhang, "Generalized fuzzy rough sets," *Inf. Sci.*, vol. 151, pp. 263–282, 2003.

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

[33] J. Zhou, C. Gao, W. Pedrycz, Z. Lai, and X. Yue, "Constrained shadowed sets and fast optimization algorithm," *Int. J. Intell. Syst.*, vol. 34, pp. 2655–2675, 2019.

[34] 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, Oct. 2013.

[35] H. T. Nguyen, W. Pedrycz, and V. Kreinovich, "On approximation of fuzzy sets by crisp sets: From continuous control-oriented defuzzication to discrete decision making," in *Proc. Int. Conf. Intell. Technologies*, 2000, pp. 254–260.

[36] P. Grzegorzewski, "Fuzzy number approximation via shadowed sets," *Inf. Sci.*, vol. 225, pp. 35–46, 2013.

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

[38] S. Mitra, H. Banka, and W. Pedrycz, "Rough fuzzy collaborative clustering," *IEEE Trans. Syst., Man, Cybern., Part B*, vol. 36, no. 4, pp. 795–805, Aug. 2006.

[39] Y. P. Zhang, F. L. Chung, and S. T. Wang, "A multiview and multiexemplar fuzzy clustering approach: Theoretical analysis and experimental studies," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 8, pp. 1543–1557, Aug. 2019.

[40] P. Lingras and C. West, "Interval set clustering of web users with rough k-means," *J. Intell. Inf. Syst.*, vol. 23, no. 1, pp. 5–16, 2004.

[41] G. Peters, "Some refinements of rough k-means clustering," *Pattern Recognit.*, vol. 39, no. 8, pp. 1481–1491, 2006.

[42] 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.

[43] P. Maji and S. Paul, "Rough fuzzy clustering for grouping functionally similar genes from microarray data," *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, vol. 10, no. 2, pp. 286–299, Mar./Apr. 2013.

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

[45] J. Zhou *et al.*, "Multigranulation rough-fuzzy clustering based on shadowed sets," *Inf. Sci.*, vol. 507, pp. 553–573, 2020.

[46] D. Ciucci, D. Dubois, and J. Lawry, "Borderline vs. unknown: Comparing three-valued representations of imperfect information," *Int. J. Approx. Reasoning*, vol. 55, pp. 1866–1889, 2014.

[47] G. M. Lang, D. Q. Miao, and H. Fujita, "Three-way group conflict analysis based on Pythagorean fuzzy set theory," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 3, pp. 447–461, Mar. 2020.

[48] Y. Li, L. Zhang, Y. Xu, Y. Yao, R. Y. K. Lau, and Y. Wu, "Enhancing binary classification by modeling uncertain boundary in three-way decisions," *IEEE Trans. Knowl.. Data Eng.*, vol. 29, no. 7, pp. 1438–1451, Jul. 2017.

[49] D. C. Liang and D. Liu, "Deriving three-way decisions from intuitionistic fuzzy decision-theoretic rough sets," *Inf. Sci.*, vol. 300, pp. 28–48, 2015.

[50] Q. H. Zhang, D. Y. Xia, K. X. 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.

[51] B. Q. Hu, "Three-way decision spaces based on partially ordered sets and three-way decisions based on hesitant fuzzy sets," *Knowl.-Based Syst.*, vol. 91, pp. 16–31, 2016.

[52] H. Y. Zhang, S. Y. Yang, and J. M. Ma, "Ranking interval sets based on inclusion measures and applications to three-way decisions," *Knowl.-Based Syst.*, vol. 91, pp. 62–70, 2016.

[53] J. T. Yao and N. Azam, "Web-based medical decision support systems for three-way medical decision making with game-theoretic rough sets," *IEEE Trans. Fuzzy Syst.*, vol. 23, no. 1, pp. 3–15, Feb. 2015.

[54] H. X. Li, L. Zhang, B. Huang, and X. Zhou, "Sequential three-way decision and granulation for cost-sensitive face recognition," *Knowl.-Based Syst.*, vol. 91, pp. 241–251, 2016.

[55] J. J. Qi, L. Wei, and Y. Y. Yao, "Three-way formal concept analysis," in *Proc. Int. Conf. Rough Sets Knowl. Technol.*, 2014, pp. 732–741.

[56] H. Yu, C. Zhang, and G. Y. Wang, "A tree-based incremental overlapping clustering method using the three-way decision theory," *Knowl.-Based Syst.*, vol. 91, pp. 189–203, 2016.

[57] P. X. Wang, H. Shi, X. B. Yang, and J. Mi, "Three-way K-means: Integrating K-means and three-way decision," *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 10, pp. 2767–2777, 2019.

[58] X. Y. Jia, W. W. Li, and L. Shang, "A multiphase cost-sensitive learning method based on the multiclass three-way decision-theoretic rough set model," *Inf. Sci.*, vol. 485, pp. 248–262, 2019.

[59] X. D. Yue, Y. F. Chen, D. Q. Miao, and H. Fujita, "Fuzzy neighborhood covering for three-way classification," *Inf. Sci.*, vol. 507, pp. 795–808, 2020.

[60] C. Luo *et al.*, "Updating three-way decisions in incomplete multi-scale information systems," *Inf. Sci.*, vol. 476, pp. 274–289, 2019.

[61] H. R. Zhang and F. Min, "Three-way recommender systems based on random forests," *Knowl.-Based Syst.*, vol. 91, pp. 275–286, 2016.

[62] B. Zhou, Y. Y. Yao, and J. G. Luo, "Cost-sensitive three-way email spam filtering," *J. Intell. Inf. Syst.*, vol. 42, no. 1, pp. 19–45, 2014.

[63] D. Liu, T. R. Li, and D. C. Liang, "Three-way government decision analysis with decision-theoretic rough sets," *Int. J. Uncertainty Fuzziness Knowl.-Based Syst.*, vol. 20, pp. 119–132, 2012.

[64] J. F. Peters and S. Ramanna, "Proximal three-way decisions: Theory and applications in social networks," *Knowl.-Based Syst.*, vol. 91, pp. 4–15, 2016.

[65] Y. B. Zhang *et al.*, "Three-way enhanced convolutional neural networks for sentence-level sentiment classification," *Inf. Sci.*, vol. 477, pp. 55–64, 2019.

[66] W. W. Li, Z. Q. Huang, and Q. Li, "Three-way decisions based software defect prediction," *Knowl.-Based Syst.*, vol. 91, pp. 263–274, 2016.

[67] D. R. Wilson and T. R. Martinez, "Improved heterogeneous distance functions," *J. Artif. Intell. Res.*, vol. 6, pp. 1–34, 1997.

[68] R. Gilad-Bachrach, A. Navot, and N. Tishby, "Margin based feature selection-theory and algorithms," in *ACM Proc. 21st Int. Conf. Mach. Learn.*, 2004, pp. 43–50.

[69] Y. Du, Q. H. Hu, P. F. Zhu, and P. J. Ma, "Rule learning for classification based on neighborhood covering reduction," *Inf. Sci.*, vol. 181, pp. 5457–5467, 2011.

[70] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*. San Mateo, CA, USA: Morgan Kaufmann, 2005.

[71] C. Elkan, "The foundations of cost-sensitive learning," in *Proc. 17th Int. Joint Conf. Artif. Intell.*, 2001, pp. 973–978.

[72] M. U. Fayyad and K. Irani, "Multi-interval discretization of continuous-valued attributes for classification learning," in *Proc. 13th Int. Joint Conf. Artif. Intell.*, 1993, pp. 1022–1027.

Xiaodong Yue (Member, IEEE) received the Ph.D. degree in pattern recognition and intelligent system from Tongji University, Shanghai, China, in 2010.

He is currently an Associate Professor with the School of Computer Engineering and Science at Shanghai University, Shanghai, China. He has published more than 60 academic papers, which include prestigious journals and international conferences. His major research interests include the theories and applications of machine learning, in which he has the special focus on soft computing.

Mr. Yue is a Member of IEEE Computational Intelligence Society, International Rough Set Society, and Chinese Association for Artificial Intelligence (CAAI) Machine Learning Technical Committee. He served as a Program Committee (PC) Member for a number of academic conferences such as Association for the Advancement of Artificial Intelligence (AAAI) and Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD), and served as an Associate Editor for *International Journal of Approximate Reasoning and Array*.

Jie Zhou (Member, IEEE) received Ph.D. degree in pattern recognition and intelligent system from Tongji University, China, in 2011.

He is currently an Associate Professor with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China. From 2010 to 2011, he was a Visiting Scholar with the Department of Electrical and Computer Engineering (ECE) at University of Alberta, Edmonton, AB, Canada. He has authored or co-authered more than 20 academic papers in these areas. His current major research interests include uncertainty analysis, pattern recognition, data mining, and intelligent systems.

Yiyu Yao is currently a Professor of Computer Science with the University of Regina, Regina, SK, Canada. He proposed a theory of three-way decision, a decision-theoretic rough set model, and a triarchic theory of granular computing. He has authored/coauthored more than 400 papers. He was 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, and web intelligence.

Duoqian Miao is a Professor with the School of Electronics and Information Engineering at Tongji University, Shanghai, China. He has published more than 180 academic papers, which include prestigious journals and international conferences, and the h-index is 39. His main research interests include rough sets, soft computing, machine learning and intelligent systems.

Prof. Miao is a Fellow of International Rough Set Society and a director of Chinese Association for Artificial Intelligence (CAAI).