Information Sciences 607 (2022) 1372–1390

Contents lists available at ScienceDirect

Information Sciences

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

Semi-supervised shadowed sets for three-way classification on partial labeled data

X.D. Yue [a,b,*], S.W. Liu [a], Q. Qian [a], D.Q. Miao [c], C. Gao [d]

[a] School of Computer Engineering and Science, Shanghai University, China
[b] Artificial Intelligence Institute of Shanghai University, China
[c] College of Electronic and Information Engineering, Tongji University, Shanghai, China
[d] College of Computer Science and Software Engineering, Shenzhen University, Guangdong, China

ARTICLE INFO

Article history:
Received 20 September 2021
Received in revised form 15 April 2022
Accepted 13 June 2022
Available online 20 June 2022

Keywords:
Semi-supervised shadowed sets
Three-way classification
Partial labeled data

ABSTRACT

Shadowed set divides a fuzzy set into three regions through fuzzy-rough transformation to denote acceptance, rejection and uncertain decision. Based on the tri-partition property, shadowed sets are utilized to implement the machine learning methods for uncertain data analysis. The extant uncertain machine learning methods with shadowed sets include the unsupervised clustering on only unlabeled data and the supervised classification on only labeled data. However, for the partial labeled data containing both labeled and unlabeled data instances, the studies of uncertain learning methods with shadowed sets are very limited. Aiming at the requirement, in this paper, we propose a novel semi-supervised shadowed set on partial labeled data and thereby construct semi-supervised shadowed neighborhoods to implement the three-way classification of uncertain data. To construct the semi-supervised shadowed set, we reformulate the objective function of shadowed sets, in which the membership loss in fuzzy-rough transformation is weighted by labeled and unlabeled data. We also analyze the influence of labeled data to the shadowed set construction. Experiments validate that the proposed three-way classification method with semi-supervised shadowed sets is effective to utilize partial labeled data to achieve low-risk uncertain data classification.

1. Introduction

As an important paradigm of granular computing [1,2], Shadowed Sets provide an effective tool to model and process the data with uncertainty [3]. Based on the fuzzy–rough transformation, shadowed sets are constructed through mapping fuzzy memberships into a triplet set {0, [0, 1], 1}. 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]) [4,5]. As to the superiority of uncertain data analysis, shadowed sets have been widely applied in data mining [6,7], decision support systems [8,9] and image analysis [10,11].

By exploring a common tri-partitioning methodology of various kinds of soft computing models including fuzzy sets, rough sets, interval sets, many-valued logic, etc., Yao proposed Three-Way Decision (3WD) theory to construct a common framework of uncertain decision making [12]. According to the 3WD theory, the decision domain will be divided into pos-

[*] Corresponding author at: School of Computer Engineering and Science, Shanghai University, China.
E-mail address: yswantfly@shu.edu.cn (X.D. Yue).

https://doi.org/10.1016/j.ins.2022.06.062
0020-0255/© 2022 Elsevier Inc. All rights reserved.

itive, negative and boundary regions, which denote the ternary options of acceptance, rejection, and uncertain case respectively [13]. In general, shadowed sets can be considered as a three-way approximation of fuzzy sets through the fuzzy-rough transformation [14]. Therefore, it is natural to utilize shadowed sets to implement three-way decision models. Based on the principles of minimum distance and least decision cost, Yao proposed an optimization-based framework to construct three-way approximations of fuzzy sets to implement the shadowed sets for uncertain decision making [5]. Zhou further investigated the mathematical properties of Yao's optimization objective function and thereby proposed the constraints for three-way approximation of fuzzy sets and implemented a constructive algorithm to generate the optimal membership threshold of shadowed sets [15]. Yao and Zhang proposed a game-theoretic approach to form shadowed sets from the three-way trade-off perspective based on game theory [16]. Campagner and Zhang proposed the strategies to compute the optimal threshold for tri-partitioning fuzzy memberships to form shadowed sets based on information entropies [17,18].

Through combining with machine learning methods, shadowed sets have been used to implement three-way clustering and classification methods for uncertain data analysis. Based on the fuzzy-rough transformation of fuzzy memberships of clusters, shadowed clustering was proposed as a uniform framework to bridge between fuzzy clustering [19,20] and rough clustering [21]. Mitra proposed a shadowed C-means algorithm that integrates fuzzy and rough clustering [22]. Zhou proposed a rough-fuzzy clustering method based on shadowed sets, in which the certain and uncertain regions of clusters are determined through optimizing the shadow thresholds [23]. For the uncertain data classification, Yue extended the neighborhoods with shadowed sets to model uncertain data and thereby designed a three-way classification method based on the shadowed neighborhoods [24]. Moreover, a strategy for accelerating shadowed set construction was proposed to improve the efficiency of the classification methods based on shadowed sets [25].

Although shadowed sets have been widely used in machine learning for uncertain data analysis, the existing studies focus on either unsupervised shadowed clustering of unlabeled data or supervised uncertain classification of labeled data. For particular classification tasks, e.g. medical image classifications, it is difficult to obtain sufficient labeled data for training classifiers [26,27]. The strategy of semi-supervised learning is required to combine limited labeled data with a large amount of unlabeled data to build up the classifiers [28,29]. The semi-supervised classification methods can utilize both labeled and unlabeled data to improve the classification performances and in the meantime reduce the dependence of data labeling [30]. However, for the partial labeled data, the semi-supervised uncertain classification method based on shadowed sets is still lacked. Aiming to overcome the shortage, in this paper, we propose a novel semi-supervised shadowed set on partial labeled data and construct semi-supervised shadowed neighborhoods to implement three-way classification of uncertain data. The contributions of this article are summarized as follows.

1. *Propose semi-supervised shadowed sets on partial labeled data.* We construct the objective function of semi-supervised shadowed sets, in which the membership loss of fuzzy-rough transformation is weighted by labeled and unlabeled data. Through minimizing the objective function of membership loss involving partial class label information, we obtain the optimal membership threshold to keep the shadowed set formulation consistent with the labeled data.
2. *Utilize semi-supervised shadowed sets to construct semi-supervised shadowed neighborhoods.* We initially build up semi-supervised fuzzy neighborhoods with semi-supervised fuzzy clustering. Based on this, we transform the fuzzy neighborhoods to semi-supervised shadowed neighborhoods through constructing the semi-supervised shadowed sets on fuzzy neighborhood memberships. The partition of certain and uncertain regions in the semi-supervised shadowed neighborhoods will be influenced by partial labeled data.
3. *Implement a three-way classification method for partial labeled data based on semi-supervised shadowed neighborhoods.* For an unknown data instance, according to its memberships to neighborhoods, we can determine the region location of the data instance respect to all shadowed neighborhoods. Considering different situations of instance locations, we design the groups of three-way classification rules within and beyond neighborhoods respectively to classify the data instance into a certain class or uncertain case.

The remainder of this paper is organized as follows. Section 2 introduces the related works of shadowed sets and three-way decisions. In Section 3, we propose a novel semi-supervised shadowed set and also analyze the influence caused by partial labeled data to the construction of shadowed sets. Section 4 introduces the method of constructing the semi-supervised shadowed neighborhoods and the corresponding three-way classification algorithm. In Section 5, experimental results validate that the proposed three-way classification method with semi-supervised shadowed neighborhoods is effective to handle partial labeled data. The work conclusion is given in Section 6.

2. Related work

2.1. Foundation of shadowed sets

Shadowed sets are constructed based on fuzzy sets through the fuzzy-rough transformation [31,32]. Suppose a discrete fuzzy set for a concept $F = \{(x_i, \mu_i)\}(i = 1, 2, \ldots, N)$, μ_i is the fuzzy membership value of the data instance x_i . Transforming the fuzzy set into a shadowed set, the fuzzy membership values of data instances are mapping into a triplet set $\{0, [0, 1], 1\}$ based on uncertainty variation. In the tripartition of fuzzy memberships, the low fuzzy memberships no more than the threshold 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 fuzzy memberships locating in the interval $(\alpha, 1-\alpha)$ constitute the shadow region. The uncertainty of a shadowed set is represented by the number of the uncertain instances in the shadow region and the variation of uncertainty in constructing a shadowed set is illustrated in Fig. 1. Given a fuzzy set, the shadow threshold α determines the shadowed set construction and is computed through optimizing the following objective of uncertainty variation [3], (See Fig. 2)

$$V(\alpha) = |\sum_{\mu_i \leqslant \alpha} \mu_i + \sum_{\mu_i \geqslant (1-\alpha)} (1-\mu_i) - \underset{\alpha<\mu_i<1-\alpha}{card}(\mu_i)|. \tag{1}$$

The uncertainty variation $V(\alpha)$ consists of the uncertainty decrement in the certain regions and the uncertainty increment in the uncertain shadow region. With the tradeoff between the uncertainty decrement and increment, $V(\alpha)$ can be also considered as the measure of fuzzy membership loss in shadowed set construction. The optimal threshold parameter α^* should achieve the balance between the uncertain shadow region and certain regions through minimizing the membership loss $V(\alpha)$.

For different data analysis tasks, the traditional shadowed set model has been widely extended. Yao summarized the principles to construct shadowed sets including the strategies of minimizing distance and cost [5]. Tahayori et al. represented the fuzziness of a fuzzy set as a gradual number and determined the shadow threshold in shadowed set construction through defuzzification of the gradual number [14]. Zhang and Yao applied a principle of tradeoff with games in order to determine the thresholds of three-way approximations in the shadowed set context [16]. Zhou proposed a constrained shadowed sets to implement a fast optimization algorithm to compute the thresholds for constructing shadowed sets [25]. Zhang et al. proposed a interval shadowed set model based on fuzzy entropy [7] and combined the fuzzy entropy with game theory to construct fuzzy-entropy game theoretic shadowed sets from the perspective of fuzzy entropy loss [18]. Gao et al. adopted the mean entropy as the basis of uncertainty measure to construct shadowed sets [33]. To make shadowed sets suitable for uncertain data classification, we reformulated the objective function of shadowed set as follows [24].

$$V(\alpha) = \lambda \cdot \left[\sum_{\mu_i \leqslant \alpha} \mu_i + \sum_{\mu_i \geqslant (1-\alpha)} (1-\mu_i)\right] + \sum_{\alpha<\mu_i<1-\alpha} |0.5-\mu_i|. \tag{2}$$

The first part of the objective function denotes the membership loss in the certain region and the second part denotes the membership loss in uncertain region. The parameter λ is a balance factor that make a tradeoff between the two parts of membership loss.

In addition to the extension of shadowed set model, shadowed sets have been used to implement machine learning methods to handle the uncertainty in data analysis applications. Based on the tripartition structure of shadowed sets, fuzzy clustering and rough clustering were represented in a uniform framework of shadowed clustering and the thresholds for partitioning the certain and uncertain regions of clusters were determined through optimizing the shadowed sets [22,34]. In [6], four kinds of shadowed sets were constructed for linguistic word modeling based on surveyed interval data. He et al. proposed an extended TODIM method based on shadowed sets to solve large-scale group decision making problem

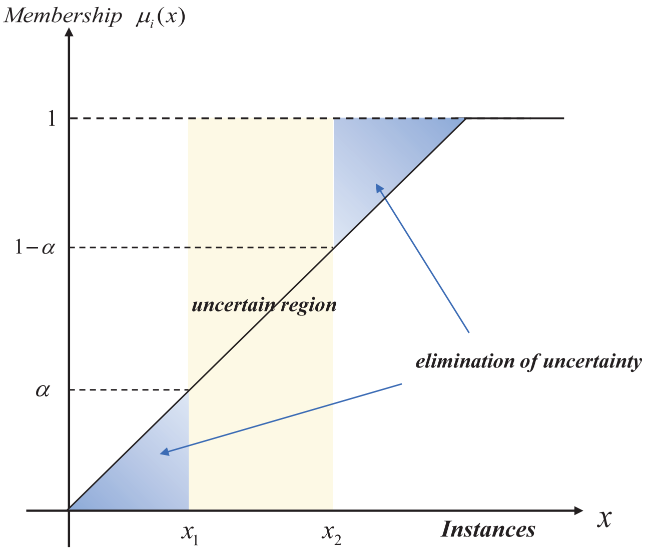


Fig. 1. Uncertainty variation in constructing a shadowed set based on a fuzzy set.

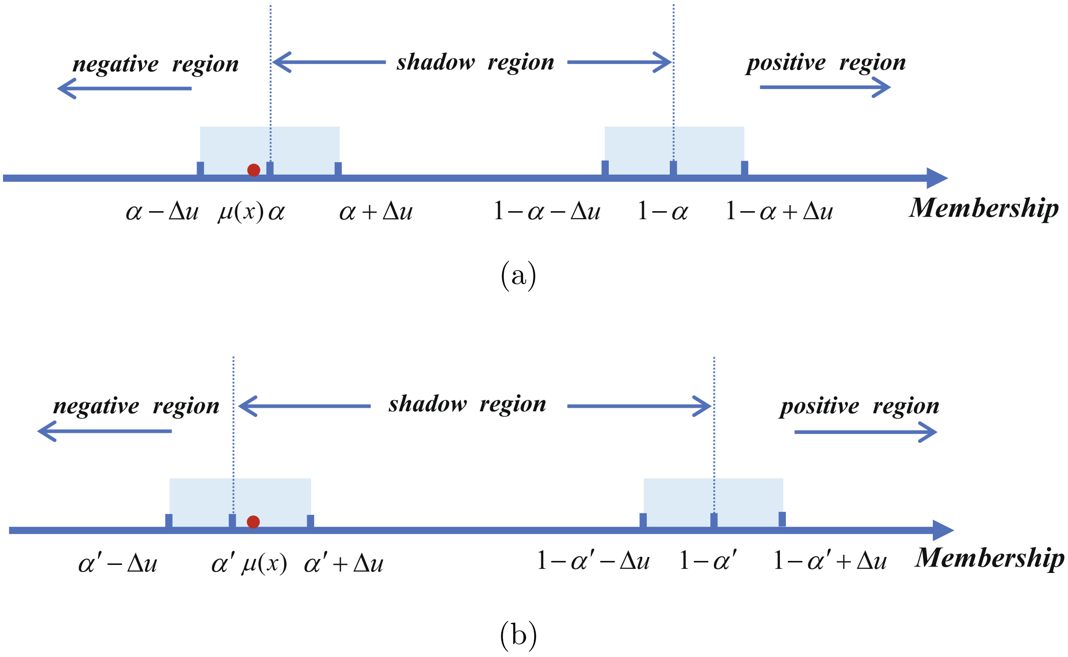


Fig. 2. Illustration of Theorem 1, (a) the original partition of shadowed set in which x of the same class as concept locates in the negative region, (b) the partition of shadowed set is adjusted to transfer x from the negative region to shadow region through minimizing membership loss.

with linguistic information [9]. In [10], a shadowed-set-based image retrieval algorithm was proposed and the shadowed sets were used to automatically determine the shadowed regions of image content for image retrieval. Moreover, shadowed sets were also used to construct recommender system for three-way recommendation, which facilitates to reduce the decision risk in recommendation [8].

In general, extant machine learning methods based on shadowed sets focused on the topics of supervised and unsupervised learning, such as the classification with full labeled data and the clustering without class labels. For partial labeled data, it is required to extend the shadowed set model to implement the semi-supervised learning method to handle the uncertain data.

2.2. Methodology of three-way decisions

Three-Way Decision methodology (3WD) is an extension of binary-decision model through adding a third option [12]. A universal set will be divided into Positive region, Negative region and Boundary region, which denote the regions of acceptance, rejection and non-commitment for ternary classification [13]. The three-way decision models are formulated through thresholding the ordered evaluation of acceptance and are implemented based on the common tri-partitioning property of many soft computing models, such as Interval Sets, Rough Sets, Fuzzy Sets and Shadowed Sets.

Suppose $(L,\preceq)$ is an 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

$$\begin{aligned}&POS_{\alpha,\beta}(v)=\{x\in U|v(x)\succeq\alpha\},\\&NEG_{\alpha,\beta}(v)=\{x\in U|v(x)\preceq\beta\},\\&BND_{\alpha,\beta}(v)=\{x\in U|\alpha\prec v(x)\prec\beta\}\end{aligned}\tag{3}$$

By measuring the uncertainty of data in classification and abstain the instances with high uncertainty, three-way decision methodology has been combined with various kinds of machine learning methods to implement the three-way learning for uncertain data analysis, such as three-way sequential rule learning [35,36], uncertain clustering [37,38], cost-sensitive classification [39–41], three-way active learning [42] and etc. Through deferring the decision making of uncertain cases, the three-way learning methods facilitate to reduce the decision risk in decision support systems, such as the medical diagnosis systems [43,44].

Although the methodology of three-way decisions has been investigated in many areas, its applications in the semi-supervised learning on partial labeled data are still limited. Gao et al. proposed a three-way strategy to partition the unlabeled data into useful, useless and uncertain data and transferred the useful unlabeled data to the co-training model for pseudo labeling to improve the classification on partial labeled data [45]. Aiming at the limitation of three-way decisions on partial labeled data, in this paper, we directly formulate the semi-supervised shadowed set on partial labeled data and thereby construct the semi-supervised shadowed neighborhoods to implement the three-way classification method. The proposed method utilizes both numerous unlabeled data and a limited number of labeled data instances to form the certain and uncertain decision regions for classification and builds up an effective semi-supervised three-way classifier on partial labeled data.

3. Semi-supervised shadowed sets

3.1. Constructing semi-supervised shadowed sets with partially labeled data

Considering the partial labeled data in computing fuzzy memberships, we extend the shadowed sets in [24] to semi-supervised ones. The partial label information will influence the threshold values of shadowed sets and thereby change the partition of certain and uncertain shadow regions. The objective function of semi-supervised shadowed set is formulated as

$$V(\alpha) = \sum_{\mu(x)\leqslant\alpha} \omega(x)\mu(x) + \sum_{\mu(x)\geqslant 1-\alpha} \omega(x)(1-\mu(x)) + \sum_{\alpha<\mu(x)<(1-\alpha)} \omega(x)|0.5-\mu(x)|, \tag{4}$$

where $\omega(x)$ is a coefficient function and $\alpha \in [0, 0.5]$ is the partition threshold parameter of semi-supervised shadowed set. $V(\alpha)$ indicates the membership loss of constructing shadowed sets and consists of three parts, the first two parts denote the membership loss of certain regions, and the third part denotes the membership loss of uncertain region. According to the threshold α, the memberships of data instances belonging to a concept will be partitioned into three regions: the positive region represented by the enhanced membership grade 1, the negative region represented by the decreased membership grade 0, and the boundary region represented by the typical membership grade 0.5, which has the greatest uncertainty.

Different from the traditional shadowed sets constructed based on only fuzzy memberships, the semi-supervised shadowed sets are constructed based on both fuzzy memberships and class labels of data. Comparing with the objective functions of the traditional shadowed sets, see Eqs. (1) and (2), the objective function of Eq. (4) is formulated with the fuzzy membership $\mu(x)$ and the coefficient function $\omega(x)$ that contains the class label information. The coefficient function $\omega(x)$ is defined as

$$\omega(x) = 1 + I(x) * \Psi(x), \tag{5}$$

where $I(x)$ is an indicator function to indicate whether the data instance x is labeled or not and whether the class label of the instance is consistent with the class of concept,

$$I(x) = \begin{cases} 0, & \textit{if} \quad x \textit{ is unlabeled}, \\ 1, & \textit{label}(x) = \textit{label}(\textit{concept}), \\ -1, & \textit{label}(x) \neq \textit{label}(\textit{concept}). \end{cases} \tag{6}$$

$\Psi(x)$ is the penalty function to denote the degree of the influence caused by the labeled data in different regions of shadowed sets,

$$\Psi(x) = \begin{cases} -\lambda, \mu(x) \geqslant 1-\alpha, \\ \frac{1}{2}\lambda, \alpha < \mu(x) < 1-\alpha, \\ \lambda, \mu(x) \leqslant \alpha. \end{cases} \tag{7}$$

The penalty factor λ is a positive constant. Without considering the class label of data, $\forall x, I(x) = 0, \omega(x) \equiv 1$, and the objective function of the semi-supervised shadowed sets returns to

$$V(\alpha) = \sum_{\mu(x)\leqslant\alpha} \mu(x) + \sum_{\mu(x)\geqslant 1-\alpha} (1-\mu(x)) + \sum_{\alpha<\mu(x)<(1-\alpha)} |0.5-\mu(x)|, \tag{8}$$

which is same as the objective function of the traditional shadowed set presented by Eq. (2).

With the objective function containing fuzzy memberships and class labels, the construction of semi-supervised shadowed sets utilizes both the unlabeled data and labeled data. Next we provide an example to further interpret the differences between the objective functions of the traditional shadowed sets and the proposed semi-supervised one.

Case study. Given a fuzzy set $F = \left\{\frac{0.1}{x_1}, \frac{0.3}{x_2}, \frac{0.5}{x_3}, \frac{0.7}{x_4}, \frac{0.9}{x_5}\right\}$ of a concept C and setting the threshold $\alpha = 0.2$, we can partition the fuzzy set into three subsets $F_N = \left\{\frac{0.1}{x_1}\right\}, F_S = \left\{\frac{0.3}{x_2}, \frac{0.5}{x_3}, \frac{0.7}{x_4}\right\}, F_P = \left\{\frac{0.9}{x_5}\right\}$ to denote the negative region, shadow region and positive region respectively. Suppose the data instances $\{x_1, x_2, x_3, x_4, x_5\}$ are partial labeled, we only know the class label of x_5 and

$class(x_5) \neq class(C)$. Setting $\lambda = 1$, according to the objective functions of different shadowed sets, we can compute the membership losses as follows.

- Eq. (1) of the shadowed set in [3]:

$$(V_1 = |0.1 + (1 - 0.9) - card\{F_S\}| = |0.1 + 0.1 - 3| = 2.8.$$

- Eq. (2) of the shadowed set in [24]:

$$V_2 = 0.1 + (1 - 0.9) + |0.5 - 0.3| + |0.5 - 0.5| + |0.5 - 0.7| = 0.6.$$

- Eq. (4) of the semi-supervised shadowed set:

$$\begin{aligned} &I(x_1) = I(x_2) = I(x_3) = I(x_4) = 0, \quad I(x_5) = -1, \Psi(x_2) = \Psi(x_3) = \Psi(x_4) = 0.5, \quad \Psi(x_1) = 1, \quad \Psi(x_5) = -1, \omega(x_1) \\ &\quad = \omega(x_2) = \omega(x_3) = \omega(x_4) = 1, \quad \omega(x_5) = 1 + (-1) \times (-1) = 2, V_3 \\ &\quad = 1 \times 0.1 + 2 \times (1 - 0.9) + 1 \times (|0.5 - 0.3| + |0.5 - 0.5| + |0.5 - 0.7|) = 0.7. \end{aligned}$$

We can find that the measure of fuzzy membership loss induced by V_1 is much coarser than V_2, V_3. Moreover, V_1 and V_2 of the traditional shadowed sets depend on only fuzzy memberships without considering class labels, which make the shadowed set construction independent to the labeled data. In contrast, the objective function V_3 of semi-supervised shadowed set is computed based on both memberships and coefficient function of class label and thereby involves the class information of partial labeled data into the shadowed set construction. As shown in the example, if a labeled data instance having different class with the concept but is partitioned into the positive region by the threshold, comparing with V_2, the function V_3 of semi-supervised shadowed set will generate higher loss to indicate the improper region partition.

Through minimizing the Eq. 4 of fuzzy membership loss with class label information, we can obtain the optimal threshold α_{opt} to construct the semi-supervised shadowed set.

$$\alpha_{opt} = argMin(V(\alpha)). \tag{9}$$

To solve the objective function, we update the threshold α to α' and adopt the updated membership loss $V(\alpha') - V(\alpha)$ as the discrete gradient to search α_{opt} by gradient descent.

Different from the shadowed sets without considering class labels, in the construction of semi-supervised shadowed sets, the class label information will influence the computation of membership loss. For a data instance x having a class label same as the concept, $I(x) = 1$, suppose the membership loss of the single instance is $V(\alpha, x)$, we infer the following three situations.

- If the labeled data x is in the positive region of the concept, this confirms the positive region of concept certainly belonging to the class. We have $\Psi(x) = -\lambda, \omega(x) = 1 - \lambda$, and the membership loss of x will be reduced to $V(\alpha, x) = (1 - \lambda)(1 - \mu(x)) < (1 - \mu(x))$. The correct class label information reduces the loss of constructing shadowed sets.
- If the labeled data x is in the boundary region of concept, $\Psi(x) = -\frac{1}{2}\lambda$, the labeled data is considered as an uncertain case by the shadowed set induced by threshold α. The membership loss of data x induced by shadowed set is denoted as $(1 - \frac{1}{2}\lambda)|0.5 - \mu(x)| < |0.5 - \mu(x)|$, which means the labeled data in uncertain region reduces the membership loss in constructing shadowed sets.
- If the labeled data x is in the negative region of concept, x is considered certainly not belonging to the concept, which contradicts the consistency between the class label of x and the concept class. The penalty function $\Psi(x) = \lambda$ and $\omega(x) = 1 + \lambda$ lead to the membership loss $(1 + \lambda)\mu(x) > \mu(x)$. This means the data in concept negative region but having same class with the concept will increase the membership loss in constructing shadowed sets.

For a data instance x having a class label different from the concept, $I(x) = -1$, we can also analyze the membership loss of instance $V(\alpha, x)$ in the following three situations.

- When the labeled data x locates in the positive region of a concept with different class, this means the shadowed set partitions a wrong data instance into the certain positive region. In this case, the membership loss of x is $(1 + \lambda)(1 - \mu(x)) > 1 - \mu(x)$ and the incorrect class label increases the loss of constructing shadowed sets.
- When the labeled data x is in the boundary region of a concept with different class, the membership loss of x is $(1 + \frac{1}{2}\lambda)|0.5 - \mu(x)| > |0.5 - \mu(x)|$. The inconsistent class label increases the membership loss of uncertain shadow region.
- When the labeled data x is in the negative region of a concept with different class. The different class label confirms that x does not belong to the concept and reduces the membership loss of negative region $(1 - \lambda)\mu(x) < \mu(x)$.

Comparing with the traditional shadowed sets without considering the class label of data, in the semi-supervised shadowed sets, the partial labeled data will influence the calculation of weights $\omega(x)$ and thereby adjust the membership loss of positive, negative and boundary region in shadowed set construction. According to this, we can optimize the threshold α of shadowed sets based on the class label information of data instances.

3.2. Theoretical analysis of semi-supervised shadowed sets

Different from the traditional shadowed sets that are constructed based on the data without class labels, in the semi-supervised shadowed sets, the partial data instances with class labels will influence the construction of shadowed sets. For example, when a data instance whose class label is same as the concept but locates in the negative region of the shadowed set, this indicates that the partition of the shadowed set is improper and the threshold parameter should be adjusted. Constructing a shadowed set on unlabeled data, we obtain the initial shadow threshold α through minimizing Eq. (8). Setting an data instance x as a labeled one in the shadowed set, x will adjust the initial shadow threshold and influence the construction of semi-supervised shadowed set.

The following theorems demonstrate the influences caused by a labeled data instance to the region partition of shadowed sets. Assume the labeled data instance will update the shadow threshold α to α', we will infer that changing the threshold in a range will reduce the loss of shadowed set construction, i.e. $V(\alpha) > V(\alpha')$ to prove the influence. According to the class label and the region location of the labeled data instance, we analyze the influences in four scenarios.

Theorem 1. *For a labeled data instance x having the same class label as the concept, when x locates in the negative region $\mu(x) \in [\alpha - \triangle u, \alpha]$, $\triangle u$ is a small positive number, if the initial shadow threshold $\alpha > \frac{2+\lambda}{8+6\lambda}$, the threshold will be reduced to enlarge the shadow region and include x into uncertain shadow region to minimize the membership loss $V(\alpha)$.*

Proof. Suppose the initial threshold of shadowed set is α, the updated threshold to transfer x from negative region into shadow region is α' and $\alpha' < \mu(x) \leqslant \alpha$, according to (4), we have

$$\begin{aligned} V(\alpha) - V(\alpha') &= (1+\lambda)\mu(x) - \left(1+\frac{\lambda}{2}\right)(0.5-\mu(x)) \\ &= (2+\tfrac{3\lambda}{2})\mu(x) - \tfrac{2+\lambda}{4} \\ &= (2+\tfrac{3\lambda}{2})(\alpha - \triangle u) - \tfrac{2+\lambda}{4} \\ &\approx (2+\tfrac{3\lambda}{2})\alpha - \tfrac{2+\lambda}{4}. \end{aligned} \tag{10}$$

If $\alpha > \frac{2+\lambda}{8+6\lambda}$, we infer that

$$V(\alpha) - V(\alpha') > \left(2+\frac{3\lambda}{2}\right)\left(\frac{2+\lambda}{8+6\lambda}\right) - \frac{2+\lambda}{4}, \left(2+\frac{3\lambda}{2}\right)\left(\frac{2+\lambda}{8+6\lambda}\right) - \frac{2+\lambda}{4} = \left(\frac{8+6\lambda}{4}\right)\left(\frac{2+\lambda}{8+6\lambda}\right) - \frac{2+\lambda}{4} = 0. \tag{11}$$

Thus we obtain $V(\alpha) - V(\alpha') > 0$ and prove that α will be reduced to α' to minimize the membership loss $V(\alpha)$.

Theorem 2. *For a labeled data instance x having the same class label as the concept, when x locates in the shadow region $\mu(x) \in [1-\alpha-\triangle u, 1-\alpha]$, $\triangle u$ is a small positive number, if the initial shadow threshold $\alpha < \frac{2+\lambda}{8-2\lambda}$, the threshold will increase to shrink the shadow region and transfer x into certain positive region to minimize the membership loss $V(\alpha)$.*

Proof. Suppose the threshold of shadowed set is α, the increased threshold to transfer x from shadow region into positive region is $\alpha', \alpha' > \alpha$ and $1-\alpha' < \mu(x) \leqslant 1-\alpha$, according to (4), we have

$$\begin{aligned} V(\alpha) - V(\alpha') &= \left(1+\frac{\lambda}{2}\right)(\mu(x)-0.5) - (1-\lambda)(1-\mu(x))) \\ &= (2-\tfrac{\lambda}{2})\mu(x) + (\tfrac{3\lambda}{4} - \tfrac{3}{2}) \\ &= (2-\tfrac{\lambda}{2})(1-\alpha-\triangle u) + (\tfrac{3\lambda}{4} - \tfrac{3}{2}) \\ &\approx (2-\tfrac{\lambda}{2})(1-\alpha) + (\tfrac{3\lambda}{4} - \tfrac{3}{2}). \end{aligned} \tag{12}$$

If $\alpha < \frac{2+\lambda}{8-2\lambda}$, we infer that

$$\begin{aligned} V(\alpha) - V(\alpha') &> \left(2-\frac{\lambda}{2}\right)\left(1-\frac{2+\lambda}{8-2\lambda}\right) + \left(\frac{3\lambda}{4} - \frac{3}{2}\right), \left(2-\frac{\lambda}{2}\right)\left(1-\frac{2+\lambda}{8-2\lambda}\right) + \left(\frac{3\lambda}{4} - \frac{3}{2}\right) \\ &= \left(\frac{8-2\lambda}{4}\right)\left(\frac{6-3\lambda}{8-2\lambda}\right) + \left(\frac{3\lambda-6}{4}\right) = 0. \end{aligned} \tag{13}$$

Thus we have $V(\alpha) - V(\alpha') > 0$ and prove that α will increase to α' to minimize the membership loss $V(\alpha)$.

Theorem 3. *For a labeled data instance x having the class label different from the concept, when x is in the shadow region $\mu(x) \in [\alpha, \alpha + \triangle u]$, $\triangle u$ is a small positive number, if the initial shadow threshold $\alpha > \frac{2-\lambda}{8-6\lambda}$, the threshold α will increase to shrink the shadow region and transfer x into negative region to minimize the membership loss $V(\alpha)$.*

Proof. Suppose the threshold of shadowed set is α, the increased threshold to transfer x from shadow region into negative region is $\alpha', \alpha' > \alpha$ and $\alpha \leqslant \mu(x) < \alpha'$, according to (4), we have

$$\begin{aligned} V(\alpha) - V(\alpha') &= \left(1 - \frac{\lambda}{2}\right)(0.5 - \mu(x)) - (1-\lambda)\mu(x) \\ &= \left(\tfrac{3\lambda}{2} - 2\right)\mu(x) + \tfrac{2-\lambda}{4} \\ &= \left(\tfrac{3\lambda}{2} - 2\right)(\alpha + \triangle u) + \tfrac{2-\lambda}{4} \\ &\approx \left(\tfrac{3\lambda}{2} - 2\right)(\alpha) + \tfrac{2-\lambda}{4}. \end{aligned} \tag{14}$$

If $\alpha > \frac{2-\lambda}{8-6\lambda}$, we infer that

$$V(\alpha) - V(\alpha') > \left(\frac{3\lambda}{2} - 2\right)\left(\frac{2-\lambda}{8-6\lambda}\right) + \frac{2-\lambda}{4}\left(\frac{3\lambda}{2} - 2\right)\left(\frac{2-\lambda}{8-6\lambda}\right) + \frac{2-\lambda}{4} = \left(\frac{6\lambda-8}{4}\right)\left(\frac{2-\lambda}{8-6\lambda}\right) + \frac{2-\lambda}{4} = 0. \tag{15}$$

Thus $V(\alpha) - V(\alpha') > 0$ and we prove that α will increase to α' to minimize the membership loss $V(\alpha)$.

Theorem 4. *For a labeled data instance x having the class label different from the concept, when x is in the positive region $\mu(x) \in [1-\alpha, 1-\alpha + \triangle u]$, $\triangle u$ is a small positive number, if the initial shadow threshold $\alpha > \frac{2-\lambda}{8+2\lambda}$, the threshold will decrease to enlarge the shadow region and include x into shadow region to minimize the membership loss $V(\alpha)$.*

Proof. Suppose the threshold of shadowed set is α, the reduced threshold to transfer x from the positive region into shadow is $\alpha', \alpha' < \alpha - \triangle u$and $\alpha' < \mu(x) \leqslant 1 - \alpha + \triangle u < 1 - \alpha'$, according to (4), we have

$$\begin{aligned} V(\alpha) - V(\alpha') &= (1+\lambda)(1-\mu(x)) - \left(1 - \frac{\lambda}{2}\right)(\mu(x) - 0.5) \\ &= \left(\tfrac{3\lambda}{4} + \tfrac{3}{2}\right) - \left(2 + \tfrac{\lambda}{2}\right)\mu(x)) \\ &= \left(\tfrac{3\lambda}{4} + \tfrac{3}{2}\right) - \left(2 + \tfrac{\lambda}{2}\right)(1 - \alpha + \triangle u) \\ &\approx \left(\tfrac{3\lambda}{4} + \tfrac{3}{2}\right) - \left(2 + \tfrac{\lambda}{2}\right)(1 - \alpha). \end{aligned} \tag{16}$$

If $\alpha > \frac{2-\lambda}{8+2\lambda}$, we infer that

$$\begin{aligned} V(\alpha) - V(\alpha') &> \left(\frac{3\lambda}{4} + \frac{3}{2}\right) - \left(2 + \frac{\lambda}{2}\right)\left(1 - \frac{2-\lambda}{8+2\lambda}\right)\left(\frac{3\lambda}{4} + \frac{3}{2}\right) - \left(2 + \frac{\lambda}{2}\right)\left(1 - \frac{2-\lambda}{8+2\lambda}\right) \\ &= \left(\frac{3\lambda+6}{4}\right) - \left(\frac{8+2\lambda}{4}\right)\left(\frac{6+3\lambda}{8+2\lambda}\right) = 0. \end{aligned} \tag{17}$$

We obtain that $V(\alpha) - V(\alpha') > 0$ and α will decrease to minimize $V(\alpha)$.

From the theorems above, we know that comparing with the traditional shadowed set without class label information, in the semi-supervised shadowed set, the partial labeled data instances will help to further optimize the threshold parameter α and thereby adjust the partition of shadowed set to minimize the membership loss. The influence caused by the labeled data to the shadowed set construction is illustrated in Fig. 3. Referring to the theorems, under specific conditions, adding a labeled data instance will update the region partition of the shadowed set constructed on unlabeled data.

4. Semi-supervised shadowed neighborhoods for three-way classification

Based on the semi-supervised shadowed sets, we can construct semi-supervised shadowed neighborhoods to implement the three-way classification on partial labeled data. The workflow is similar to the uncertain classification with shadowed neighborhoods [24]. First, we build up semi-supervised fuzzy neighborhoods by semi-supervised fuzzy clustering. Second, we extend the fuzzy neighborhoods to semi-supervised shadowed neighborhoods through constructing semi-supervised shadowed sets on the fuzzy memberships of neighborhoods. Finally, we design the three-way classification method with semi-supervised shadowed neighborhoods.

4.1. Semi-supervised fuzzy neighborhoods

For the partial labeled data, referring to [46], we utilize a semi-supervised fuzzy clustering method to build up semi-supervised fuzzy neighborhoods. Assuming the number of data instance is N and the data set can be partitioned into K clus-

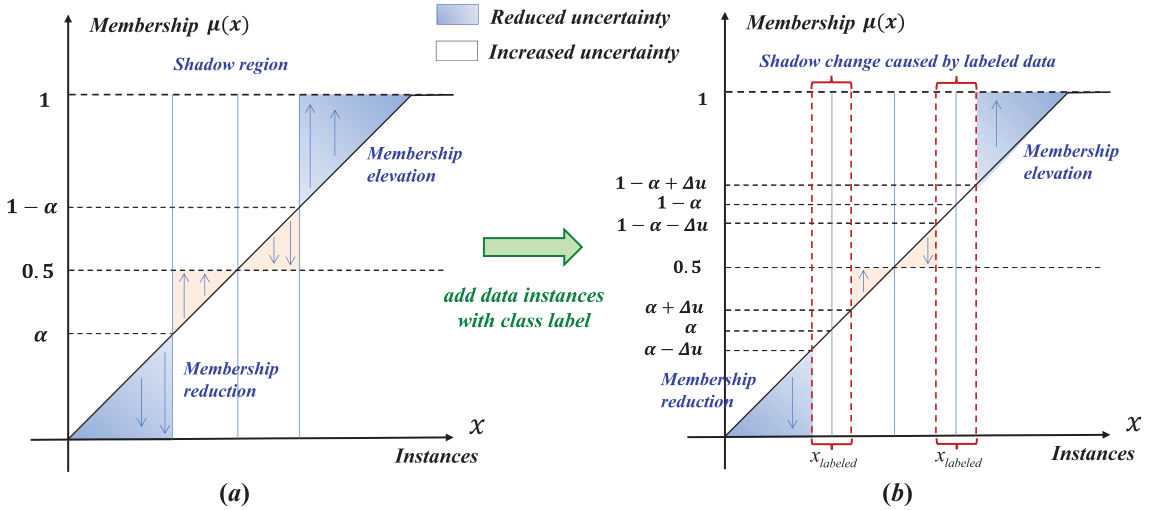


Fig. 3. The influence caused by the labeled instances to shadowed set construction. (a) Positive region, negative region and shadow region of a shadowed set, (b) updated region partition of the shadowed set caused by labeled data.

ters, $\mu_k(x_i)$ indicates the membership grade of a data instance x_i belonging to cluster k, v_k represents the prototype associated with cluster k. The objective function of semi-supervised fuzzy clustering is defined as

$$J(U,V)=\sum_{k=1}^{K}\sum_{i=1}^{N}\mu_k(x_i)^m\|x_i-v_k\|^2+\rho\sum_{k=1}^{K}\sum_{i=1}^{N}\left(\mu_k(x_i)-\tilde{\mu}_k(x_i)\right)^m\|x_i-v_k\|^2. \tag{18}$$

In the function, the first term is the objective of fuzzy C-means, the second one denotes the fuzzy cluster membership loss with respect to the class labels of partial data, ρ is the factor to control the influence of labeled data. $\tilde{\mu}_k(x_i)$ in the second term is the element of the matrix $\widetilde{U}$ and are iteratively computed as follows.

$$\tilde{\mu}_k(x_i)^{(t)}=\tilde{\mu}_k(x_i)^{(t-1)}-\beta\frac{\partial Q\left(F,\widetilde{U}\right)}{\partial\tilde{\mu}_k(x_i)}, \tag{19}$$

where t is the iteration counter, β is the learning rate and

$$Q\left(F,\widetilde{U}\right)=\sum_{h=1}^{H}\sum_{i=1}^{N}\delta_i\left(f_{ih}-\sum_{k\in\pi_h}\tilde{\mu}_k(x_i)\right)^2. \tag{20}$$

$$\delta_i=\begin{cases}1, & x_i \textit{ is labeled},\\ 0, & \textit{otherwise},\end{cases} \tag{21}$$

denotes the data instance x_i is labeled or not,

$$f_{ih}=\begin{cases}1, & x_i \textit{ belongs to class } h,\\ 0, & \textit{otherwise},\end{cases} \tag{22}$$

indicates that x_i belongs to class h or not. For a cluster containing both labeled and unlabeled data, the class of the cluster is determined depending on the majority class of the labeled data in it. π_h denotes the set of clusters belonging to class h. It is natural to assume that a class can be partitioned into several clusters. Given H classes, for each class h, there are K_h clusters in it, we have

$$\sum_{h=1}^{H}K_h=K. \tag{23}$$

Utilizing the above semi-supervised fuzzy clustering method, we can generate K clusters with partial class labels. Considering the fuzzy clusters as fuzzy neighborhoods, we can obtain K semi-supervised fuzzy neighborhoods in which each neighborhood has a fuzzy membership distribution of labeled and unlabeled data instances.

4.2. Semi-supervised shadowed neighborhoods

Based on the semi-supervised fuzzy neighborhoods, we can construct semi-supervised shadowed neighborhoods through formulating the shadowed sets on the fuzzy memberships of neighborhoods, and the threshold parameter α for constructing shadowed sets is computed according to (4), (9). Referring to the shadowed neighborhoods proposed in [24], for a data instance x, its fuzzy membership belonging to the kth neighborhood $\mu_k(x)$ is thresholded and the neighborhood is partitioned into three regions to form the shadowed neighborhood as

$$\begin{aligned} &\mu_k(x) \leqslant \alpha \Rightarrow x \in NEG_k, \\ &\alpha < \mu_k(x) < 1-\alpha \Rightarrow x \in BND_k, \\ &\mu_k(x) \geqslant 1-\alpha \Rightarrow x \in POS_k, \end{aligned} \tag{24}$$

in which POS_k, NEG_k and BND_k denote the certain positive region, certain negative region, and the uncertain boundary (shadow region) of the shadowed neighborhood k respectively.

In Section 3, we analyze the influence of the labeled data to the construction of shadowed set, we can also obtain the similar analysis results in constructing shadowed neighborhoods. When the partial labeled data have the same class as the neighborhood, Fig. 4 shows the influences of the homogeneous labeled data (class label same as the neighborhood class) to the construction of shadowed neighborhood. Fig. 4a shows a shadowed neighborhood constructed based on the data without class labels. If the homogeneous labeled data locate in the negative region of shadowed neighborhood, the threshold α will be reduced to enlarge the uncertain boundary to include the homogeneous data to minimize the membership loss, which is consistent with Theorem 1. As to Theorem 2, if the homogeneous labeled data are in the boundary region of shadowed neighborhood, the threshold α will increase to reduce the uncertain boundary to transfer the labeled homogeneous data from boundary to positive region.

Similarly, Fig. 5 illustrates the influences of the heterogeneous partial labeled data (class label different from the neighborhood class) to the construction of shadowed neighborhood. As introduced in Theorem 3, if the heterogeneous labeled data locate in the boundary region of shadowed neighborhood, the threshold α will increase to reduce the uncertain boundary to exclude the heterogeneous data from the neighborhood. As to Theorem 4, if the heterogeneous labeled data are in the positive region of shadowed neighborhood, the threshold α will be reduced to enlarge the uncertain boundary to transfer the heterogeneous data from positive region to boundary to minimize the membership loss.

We summarize the workflow of constructing semi-supervised shadowed neighborhoods on partial labeled data in the following algorithm.

Algorithm 1: Constructing semi-supervised shadowed neighborhoods on partial labeled data

```
Input: Data set $X = \{X_u, X_l\}$ of $n$ data instances, in which $X_u$ denotes unlabeled data and $X_l$ is the labeled data set;
Output: Semi-supervised shadowed neighborhoods on $X$, $O = \{O_1, \ldots, O_k, \ldots, O_K\}$;
1: Utilize semi-supervised fuzzy clustering on $X$ to construct $K$ semi-supervised fuzzy neighborhoods according to (18)–(22);
2: Determine the class of each neighborhood depending on the majority class of partial labeled data in the neighborhood;
3: Compute the threshold parameter $\alpha$ of each fuzzy neighborhood according to (4), (9) and thereby tri-partition each fuzzy neighborhood to form $K$ shadowed neighborhoods $\{O_k, k = 1, \ldots, K\}$;
4: Return the generated semi-supervised shadowed neighborhoods $O_k, k = 1, \ldots, K$.
```

As presented in Algorithm 1, constructing shadowed neighborhoods mainly suffers the computational complexity caused by semi-supervised fuzzy clustering and shadow threshold optimization. Given a partial labeled data set X and $|X| = n$, the complexity of constructing K fuzzy neighborhoods based on semi-supervised clustering is $O(I_{clu} \times K \times n), I_{clu}$ is the iteration times of clustering. For extending fuzzy neighborhoods to shadowed ones, we optimize the shadow threshold α for each shadowed neighborhood, which needs $O(I_{opt} \times K)$ calculations, I_{opt} is the iteration times of threshold parameter search. Thus the computational complexity for constructing K semi-supervised shadowed neighborhoods is summarized as $O((I_{clu} \times n + I_{opt}) \times K)$.

4.3. Three-way classification with semi-supervised shadowed neighborhoods

Utilizing the semi-supervised shadowed neighborhoods, we can implement the three-way classification on partial labeled data. Given K semi-supervised shadowed neighborhoods, for an unknown data instance x, according to (24), we can obtain the following sets to describe the region location of x in the shadowed neighborhoods.

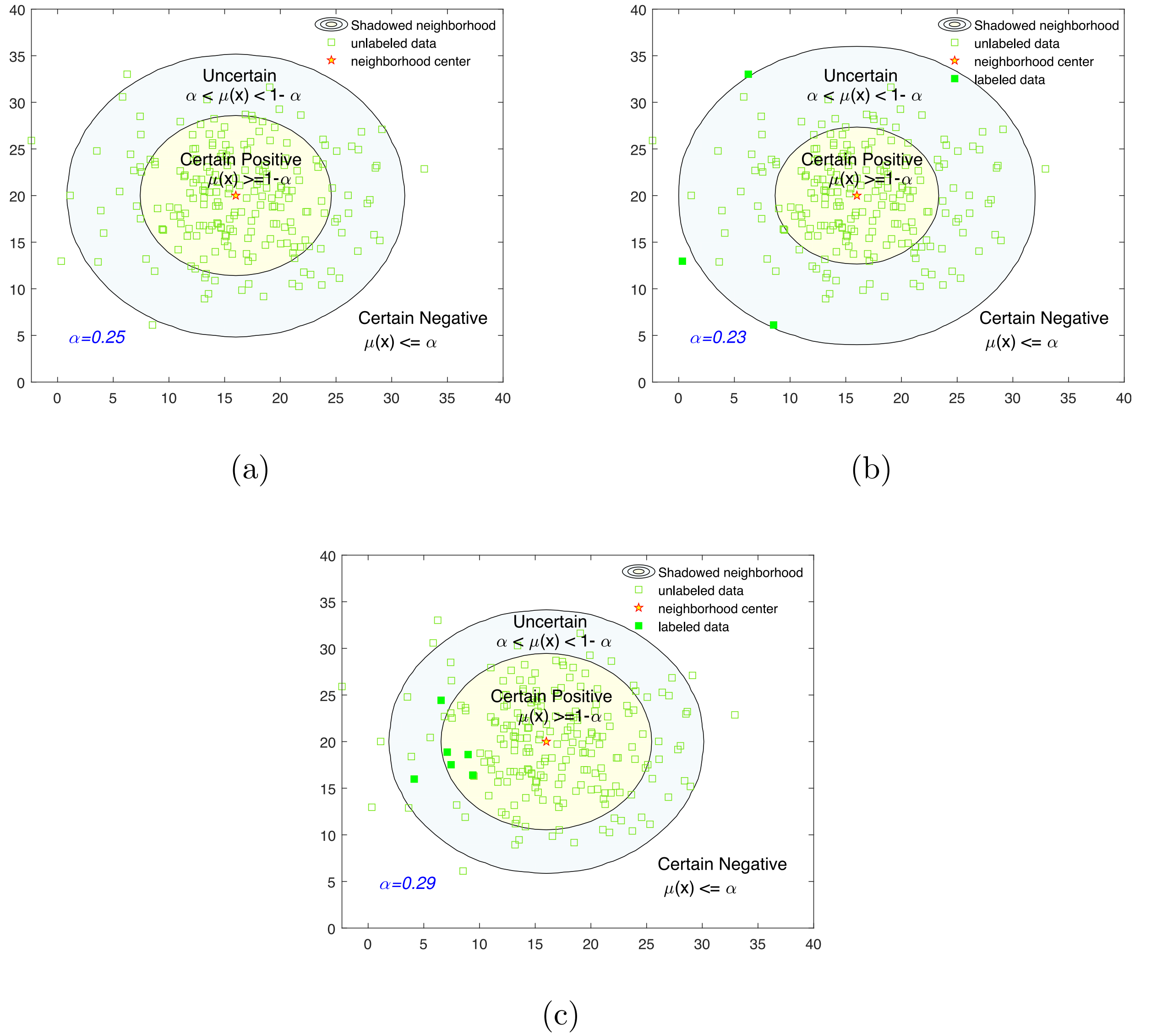


Fig. 4. (a) Shadowed neighborhood constructed on unlabeled data, (b) the homogeneous instances in negative region reduce the threshold α to enlarge uncertain boundary, (c) the homogeneous instances in boundary region increase the threshold α to shrink uncertain boundary.

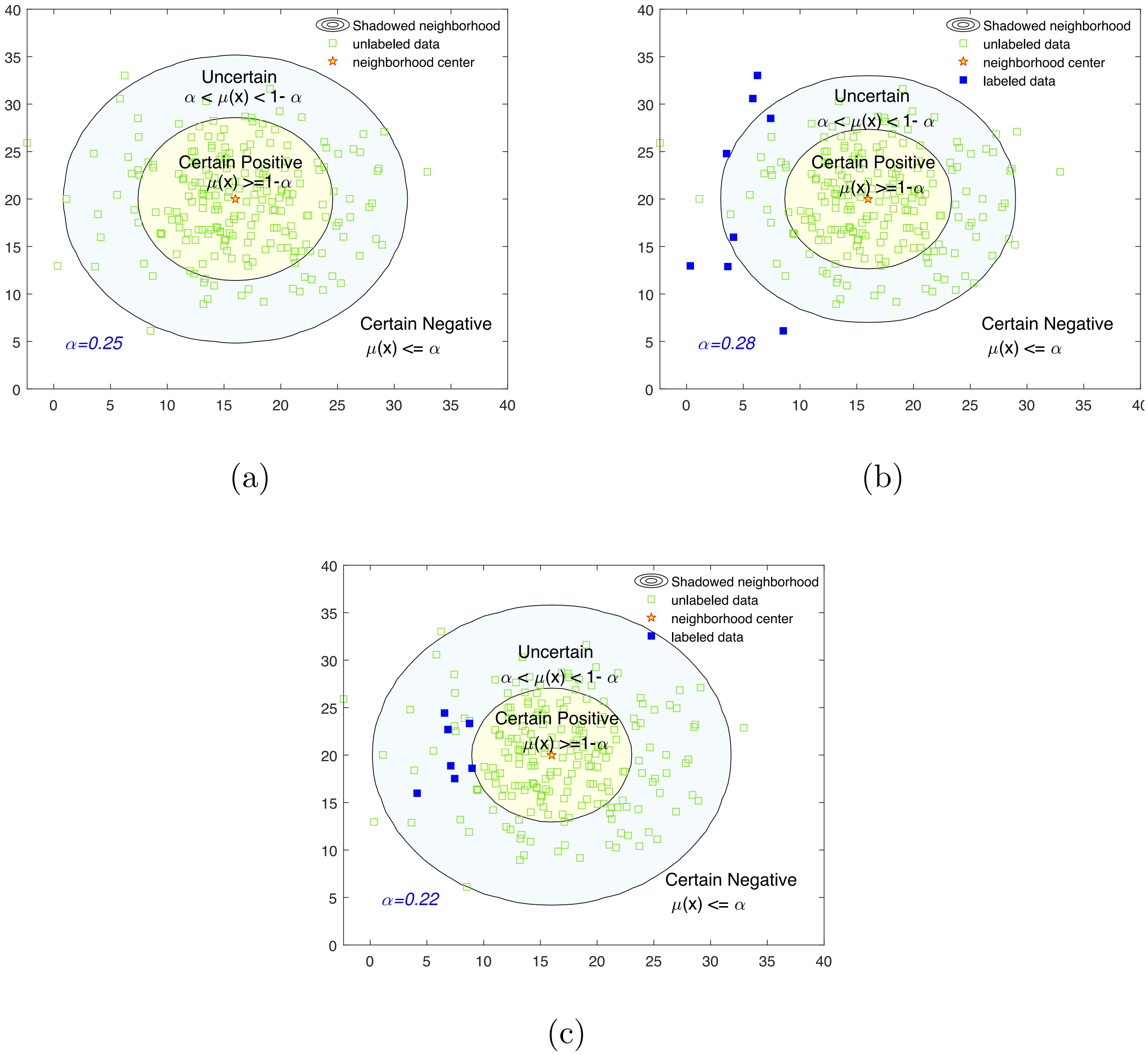
$$\begin{aligned} SNN(x) &= \{k|x \in NEG_k\}, \\ SNU(x) &= \{k|x \in BND_k\}, \\ SNP(x) &= \{k|x \in POS_k\}. \end{aligned} \tag{25}$$

Obviously, $SNP(x)$ is the set of the indexes of the shadowed neighborhoods whose positive regions containing the instance $x, SNU(x)$ is the set of the indexes of the neighborhoods in which x locates in the uncertain boundary region, and $SNN(x)$ denotes the set of the shadowed neighborhoods excluding x. Based on the locations of x in K shadowed neighborhoods $O = \{O_1, \cdots, O_k, \cdots O_K\}$, we design following three-way classification rules to classify x into certain class and uncertain class.

4.3.1. Classification rules for x within shadowed neighborhoods

Given a data instance x, if x locates within the neighborhoods of $O, \exists O_k \in O, \mu(x) > \alpha, |SNP(x)| \geqslant 1$ or $|SNU(x)| \geqslant 1$.

1. If $|SNP(x)| = 1, x$ certainly belongs to the class of the unique neighborhood in $SNP(x)$.
2. If $|SNP(x)| > 1$ and $\forall k_1, k_2 \in SNP(x), class(O_{k_1}) = class(O_{k_2}), x$ certainly belongs to the class of the neighborhoods in $SNP(x)$. Otherwise if $\exists k_1, k_2 \in SNP(x)$ and $class(O_{k_1}) \neq class(O_{k_2}), x$ belongs to multiple neighborhoods of different classes with conflict and should be classified as uncertain.



(a)

(b)

(c)

Fig. 5. (a) Shadowed neighborhood constructed on unlabeled data, (b) the heterogeneous instances in negative region increase the threshold α to reduce uncertain boundary, (c) the heterogeneous instances in boundary region reduce the threshold α to enlarge uncertain boundary.

3. If $|SNP(x)| = 0$ and $|SNU(x)| > 0$, suppose the major class of the neighborhoods in $SNU(x)$ is C and $|\{k|k \in SNU(x) \wedge class(O_k) = C\}|/|SNU(x)| \geqslant 60\%$, x is recognized as the class C. Otherwise, x is judged as an uncertain case.

The classification rules within shadowed neighborhoods indicate when x belongs to the positive regions of homogeneous shadowed neighborhoods, we can certainly classify x into the same class as the neighborhoods. Otherwise if x certainly belongs to multiple heterogenous neighborhoods, which leads to classification conflicts, x should be considered as an uncertain case. If x locates in the boundary regions of multiple neighborhoods, we determine whether x belongs to a certain class or an uncertain case depending on the majority class of neighborhoods.

4.3.2. Classification rules for x beyond shadowed neighborhoods

If a data instance x is beyond the neighborhood set $O, \forall O_k \in O, \mu_k(x) \leqslant \alpha, |SNP(x)| = 0, |SNU(x)| = 0$. We adopt two thresholds T_f, T_r of neighborhood membership to classify the uncertain data. T_f defines the minimum neighborhood membership to determine whether a data instance is far from the shadowed neighborhoods to be recognized as an uncertain (unknown) case. T_r is the threshold of neighborhood membership ratio to check whether a data instance belongs to the multiple neighborhoods of different classes, which is recognized as an uncertain case caused by conflict.

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

The classification rules beyond shadowed neighborhoods depend on the distances between instances and neighborhoods. If the instance x is too far from neighborhoods and the membership of x belonging to its nearest neighborhood is less than the threshold T_f, x is considered as an uncertain case. 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 classify x to a certain class. 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 heterogeneous neighborhoods are similar, x is considered as an uncertain case. If the membership difference exceeds 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 process of the three-way classification with semi-supervised shadowed neighborhoods is listed in the following algorithm.

Algorithm 2: Three-way classification with semi-supervised shadowed neighborhoods (3WC-SSN)

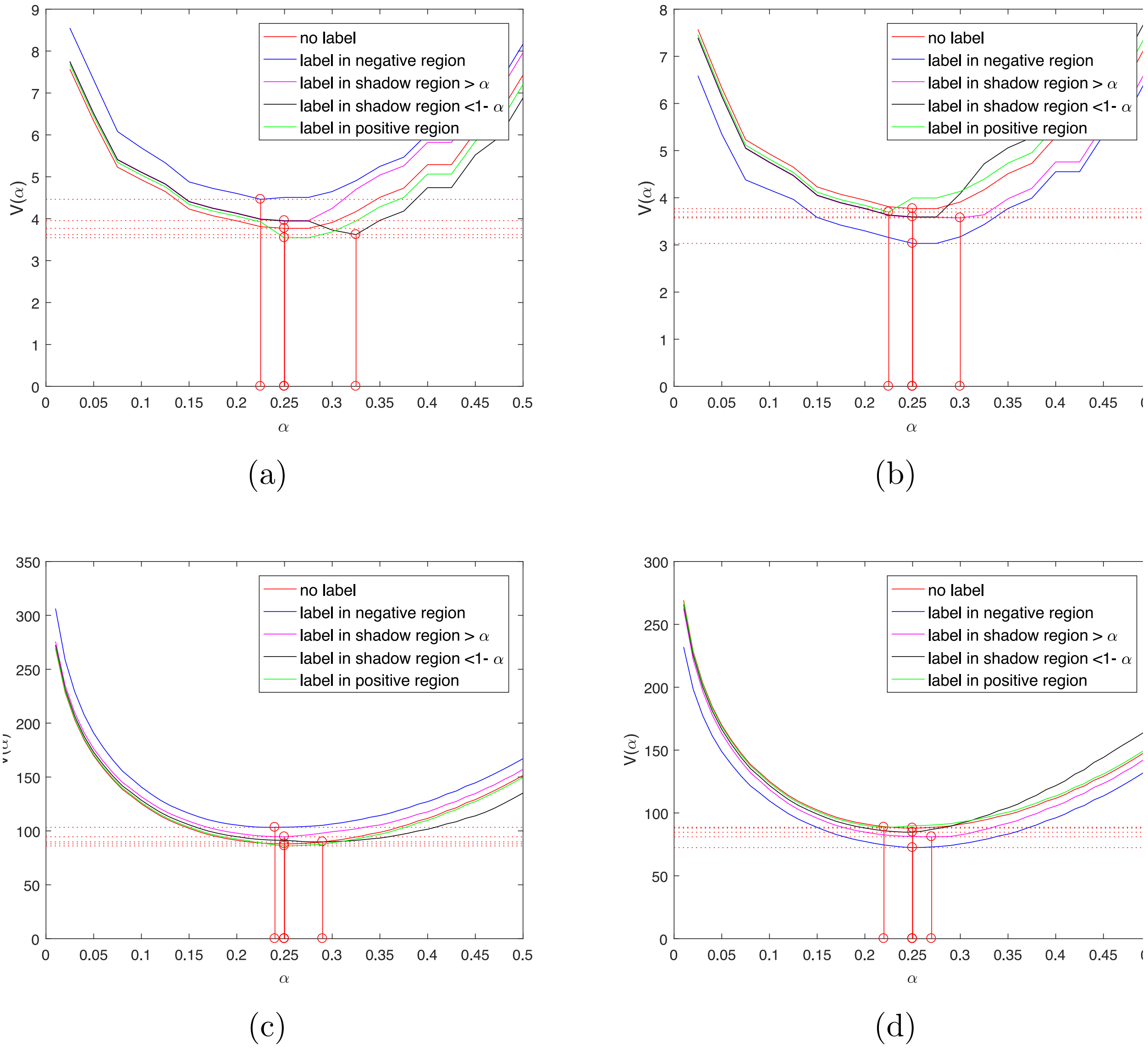
```
Input: Semi-supervised shadowed neighborhoods $O = \{O_1, O_2, \ldots, O_K\}$ constructed on the partial labeled data $X$, an unknown data instance $x$;
Output: Three-way classification result of $x$ based on $O$;
Compute the fuzzy memberships of $x$ belonging to $K$ shadowed neighborhoods and determine the region of $x$ in each shadowed neighborhood according to (24);
Generate the sets $SNP(x), SNU(x)$ and $SNN(x)$ to indicate the location of $x$ in the shadowed neighborhoods $O$ according to (25);
if $\exists O_k \in O, \mu(x) > \alpha, |SNP(x)| \geqslant 1$ or $|SNU(x)| \geqslant 1$ then
    Judge $x$ within the neighborhoods $O$ and adopt the classification rules within shadowed neighborhoods to classify $x$ into a certain class or uncertain case;
else
    if $\forall O_k \in O, \mu_k(x) \leqslant \alpha, |SNP(x)| = 0$ and $|SNU(x)| = 0$ then
        Judge $x$ beyond the neighborhoods $O$ and adopt the classification rules beyond shadowed neighborhoods to classify $x$ into a certain class or uncertain case;
    end if
end if
Return the classification result of $x$.
```

As shown in Algorithm 2, performing three-way classification for a single instance x requires $O(K)$ calculations to compute the fuzzy memberships of x belonging to K neighborhoods and obtain the region locations of x in all shadowed neighborhoods. Moreover, determining first and second nearest neighborhoods of the instance needs sorting the neighborhoods memberships and the computational complexity is $O(K \times \log K)$. Thus the total computational complexity of three-way classification of a single instance with K shadowed neighborhoods is $O(K \times (\log K + 1))$.

5. Experimental results

Different from the traditional classification methods, the three-way classifier with semi-supervised shadowed neighborhoods is built up based on partial labeled data and classifies data instances into certain classes and uncertain cases, which facilitates to reduce the classification risk. We abbreviate the proposed three-way classification method with semi-supervised shadowed neighborhoods to SSN-3WC. In order to verify the proposed three-way classification method, we implement three tests in the experiment. In the first test, we validate that the threshold optimization in constructing semi-supervised shadowed set is effective. Second, we analyze the influence of partial labeled data to the proposed semi-supervised three-way classification method. In the final test, we compare SSN-3WC with other typical semi-supervised classification methods to validates the superiority of the proposed method.

In order to overall evaluate the three-way classification methods, we adopt the measures of *accuracy*, *precision*, *recall rate*, *F1 score*, *ratio of uncertain data (UR)* and *classification cost* as the evaluation criteria. Given a data set $X = X_c \cup X_u$ and a classifier, X_c denotes the set of data instances that are assigned by certain class labels and X_u is the set of data instances that are classified as uncertain cases. Suppose P is the number of the positive-class instances and N is the number of negative-class



(a)

(b)

(c)

(d)

Fig. 6. Membership loss variation for constructing shadowed set on (a) discrete fuzzy set in which the labeled data have the same class as the concept, (b) discrete fuzzy set with the labeled data having different class from the concept, (c) continuous fuzzy set in which the labeled data have the same class as the concept, (d) continuous fuzzy set with the labeled data having different class from the concept.

instances in X_c. In the certain classification results, TP and FP denote the numbers of true positive data instances and false positive instances, TN and FN denote the numbers of true negative instances and false negative instances respectively. The calculations of the classification measures are listed as follows.

$$\begin{aligned}
&Accuracy(\%) = (TP + TN)/(P + N),\\
&Precision(\%) = TP/(TP + FP),\\
&Recallrate(\%) = TP/P,\\
&F1score(\%) = (2 \cdot Precision \cdot Recall\ rate)/(Precision + Recall\ rate),\\
&UR(\%) = |X_u|/|X|,\\
&Cost = C_{NP} \cdot \tfrac{FP}{P+N} + C_{PN} \cdot \tfrac{FN}{P+N} + C_U \cdot UR.
\end{aligned}$$

In the cost measure, we assume the correct classification suffers no cost. C_{NP}, C_{PN}, C_U denote the costs of false-positive misclassification, false-negative misclassification, and the classification of uncertain instances respectively. Suppose the positive class is minor class (risky class), the false-negative misclassification will cause more costs than the false-positive misclassification. For example, in medical image analysis, misclassifying malignant tumors as benign will lead to more risk than judg-

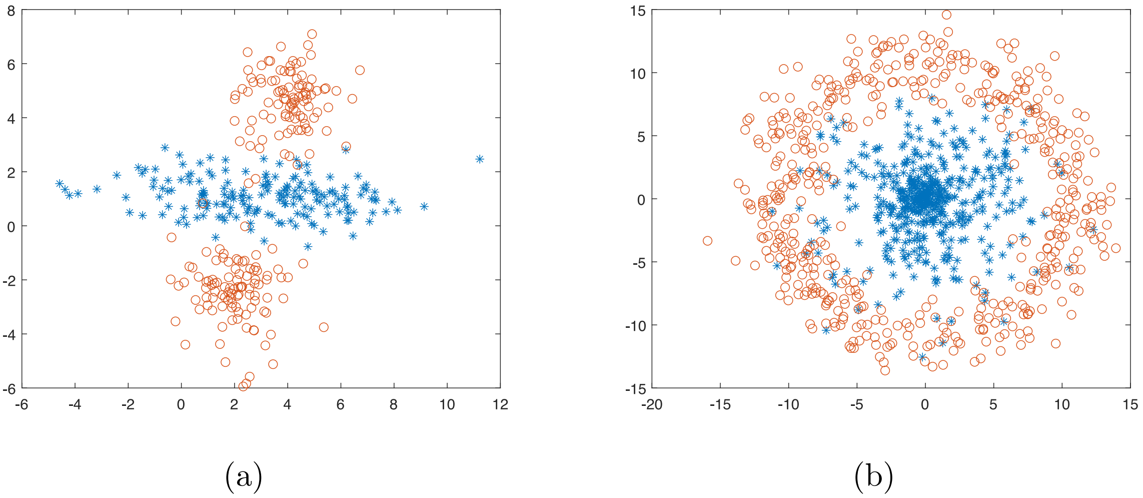


Fig. 7. Two synthetic data sets for the test of partial labeled data influence, (a) data set 1: cross-bar data, (b) data set 2: ring-shaped data.

ing benign tumor as malignant. The classification of uncertain instances will delay the decision making and lead to less cost than misclassifications. We set $C_{PN}/C_{NP}/C_U = 5/1/0.5$ in the following tests.

5.1. Test of semi-supervised shadowed set construction

In order to validate the construction of the proposed semi-supervised shadowed sets, we test the optimization of threshold parameter α in constructing the shadowed sets on partial labeled data. We generate both discrete and continuous fuzzy memberships to validate the threshold parameter optimization of shadowed sets. We randomly generate 30 fuzzy memberships to construct a discrete fuzzy set as follows.

$$F:\{\mu\} = \{0.03; 0.05; 0.07; 0.10; 0.11; 0.14; 0.17; 0.23; 0.28; 0.31; 0.33; 0.34; 0.38; 0.44; 0.45; 0.48; 0.51; \\ 0.54; 0.60; 0.64; 0.68; 0.71; 0.75; 0.78; 0.79; 0.81; 0.85; 0.93; 0.94; 0.95\}$$

Besides discrete fuzzy sets, we also adopt the following exponential function to formulate a continuous fuzzy membership function to test the threshold computation of semi-supervised shadowed sets.

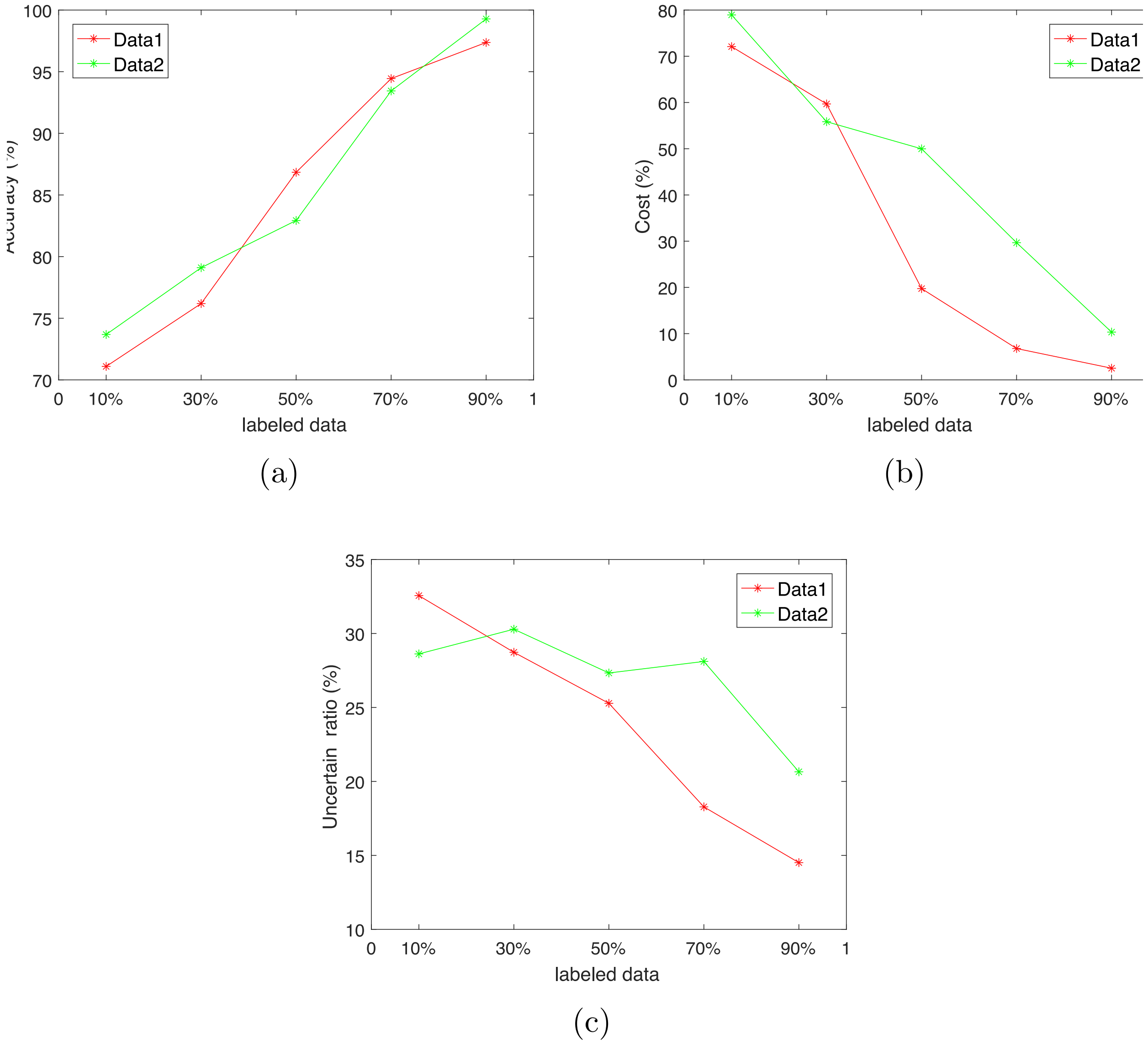
$$F:\mu(x) = e^{-\left(\frac{x-50}{20.5}\right)^2}, \tag{26}$$

in which x is valued from 0 to 100 by step 0.1.

Without considering class labels, we first initialize the shadowed set on the fuzzy set F and obtain the threshold parameter α through minimizing the objective function of (8). Second, we assign class labels to the data instances whose fuzzy memberships lie in the positive region ($\mu \geqslant 1-\alpha$), negative region ($\mu \leqslant \alpha$) and uncertain region (nearby left margin $\mu > \alpha$and right margin $\mu < 1-\alpha$) of the shadowed set respectively. Through minimizing the objective function $V(\alpha)$ of (4) with discrete gradient descent, we can obtain the optimal threshold parameters of the semi-supervised shadowed sets constructed on the partial labeled data.

Assigning class labels to a part of data instances and constructing semi-supervised shadowed sets on the discrete and continuous fuzzy sets, we illustrate the variation of membership loss $V(\alpha)$ as the threshold parameter α changing in the interval [0,0.5] in Fig. 6, in which the membership loss variations corresponding to the labeled data in different regions of shadowed sets are marked by different colors and the computed optimal thresholds are marked by red vertical lines.

When the class we assigned to the labeled data is consistent with the concept represented by the fuzzy set, Fig. 6a and 6c show the variation of membership loss $V(\alpha)$ against the threshold α. We mark the polygonal line with different colors to denote the variations of $V(\alpha)$ when labeled data instances are located in the different regions of the shadowed sets. The optimal thresholds computed via discrete gradient descent are also marked by red vertical lines. We can find that the computed thresholds generate the minimum membership loss when the labeled data instances occur in different regions of shadowed sets. Moreover, we can find that the labeled data in positive region further reduces the minimum membership loss than that without considering class labels, and the labeled data in negative region generates more membership loss than the no label case. This indicates the correct data labeling which are consistent with the concept facilitate to improve the region partition of the shadowed set and the incorrect labeling will bring about conflict and increase the costs of shadowed set construction. When the class of the labeled data is different from the concept, Fig. 6b and 6d show the variations of $V(\alpha)$ when labeled data instances are located in the different regions of the shadowed sets. In contrast to the consistent class labels, we can find that



(a)

(b)

(c)

Fig. 8. Classification results of 3WC-SSN method on two synthetic data sets with varying ratio of labeled training data, (a) accuracy variation as labeled data increasing, (b) variation of classification cost, (c) variation of uncertain instance ratio.

Table 1
Experimental data sets for semi-supervised classification.

Data sets	*Features*	*Instances*	*Class ratio*	*Attribute type*
Australian Credit	14	690	44%vs.56%	Mixed
Appendicitis	7	106	20%vs.80%	Numerical
Banknote Authentication	4	1372	44%vs.56%	Numerical
Breast Cancer Wisconsin (Original)	9	699	34%vs.66%	Numerical
Vertebral Column	7	310	32%vs.68%	Numerical
diabetes	8	768	35%vs.65%	Mixed
Ionosphere	34	351	36%vs.64%	Mixed
phoneme	6	5404	29%vs.71%	Mixed
Diabetic Retinopathy Debrecen	20	1151	47%vs.53%	Mixed
Mammographic Mass	5	961	46%vs.54%	Numerical
Wisconsin Diagnostic Breast Cancer (WDBC)	30	569	37%vs.63%	Numerical
Wisconsin Prognostic Breast Cancer (WPBC)	33	198	24%vs.76%	Numerical

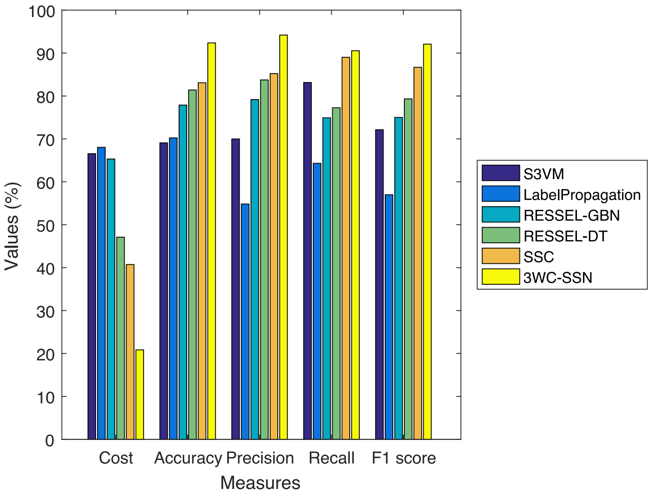


Fig. 9. Comparison of classification results generated by different semi-supervised classification methods.

Table 2
Average classification results of semi-supervised classification methods on all data sets.

Methods	*Cost* (10^{-2})	*Acc* (%)	*Prec* (%)	*Recall* (%)	*F1* (%)
S3VM	66.56	69.08	69.99	83.15	72.14
LabelPropagation	68.03	70.24	54.83	64.32	56.98
RESSEL-GBN	65.32	77.88	79.18	73.92	75.00
RESSEL-DT	47.11	81.40	83.74	77.25	79.32
SSC	40.73	83.08	85.21	89.02	86.69
3WC-SSN	**20.83**	**92.38**	**94.19**	**90.54**	**92.07**

the incorrect labeled data in negative region will reduces the membership loss and incorrect labels in positive region will increase the membership loss in constructing shadowed sets.

5.2. Test of influence of partial labeled data

In the second test, we analyze the influence of partial labeled data to the proposed semi-supervised three-way classification method 3WC-SSN. We generate two synthetic data sets of two classes to perform the test, in which the data distributions of two classes are overlapped and some instances are easily to be confused. Fig. 7 illustrates the synthetic data sets.

Ignoring the class labels of synthetic data and assign class labels to different ratios of training data from 10% to 90%, we construct the partial labeled data sets and perform the proposed 3WC-SSN method. Fig. 8 shows the classification accuracy, costs and the ratio of uncertain instances generated by 3WC-SSN method. We can find that as the number of labeled data instances increasing, the accuracy index rises and the cost is reduced, which indicates that 3WC-SSN method can make good use of the label information to improve the classification precision. Moreover, we observe that as the number of labeled data instances increasing, the ratio of uncertain instances recognized by the 3WC-SSN method is reduced, this means increasing labeled data will be helpful to confirm the classification of uncertain cases and thereby reduce the uncertainty in classification.

5.3. Comparison with other semi-supervised classification methods

In the final test, we compare the proposed 3WC-SSN method with various types of semi-supervised learning methods to validate its superiority. The comparative methods include semi-supervised support vector machine (S3VM) [29], class label propagation (LabelPropagation) [47], semi-supervised clustering for classification (SSC) [46], reliable ensemble of semi-supervised classification based on Gaussian naive Bayesian classifiers (RESSEL-GBN) and decision trees (RESSEL-DT) [48]. We perform all the semi-supervised classification methods on 12 data sets from UCI and KEEL databases. The data sets are collected from medical and economic areas and have imbalanced class distributions. We consider the minor class as the positive class (risky class) and set more costs for the false-negative misclassification. The descriptions of data sets are listed in Table 1.

Using 10-fold cross validation and employing 10% labeled data in each training data set, we obtain the average classification results on all the test data sets for each semi-supervised classification method. Fig. 9 and Table 2 present the detailed classification results. From the experimental results, we can find that comparing with the certain semi-supervised classification methods, the proposed three-way classification method based on semi-supervised shadowed sets is effective to handle uncertain data and generally achieves more precise classifications, and in the meantime produces lower classification costs. Utilizing the class information of partial labeled data, our method is able to recognize a limited number of uncertain data instances and thereby improves the classification precision and recall rate of the risky class to achieve good classification performances.

6. Conclusion

In this paper, we propose a novel semi-supervised shadowed set to construct shadowed neighborhood for three-way classification on partial labeled data. The partial labeled data will adjust the optimization of the shadow thresholds and thereby influence the construction of the shadowed sets. Constructing semi-supervised shadowed sets on fuzzy neighborhood memberships, we can formulate semi-supervised shadowed neighborhoods of the certain positive region and uncertain boundary region to involve both labeled and unlabeled data. Based on the semi-supervised shadowed neighborhoods, we design three-way classification rules to implement a three-way classification algorithm to distinguish data instances into certain classes and uncertain cases. In the experiments, through comparing with other types of semi-supervised classification methods, the proposed three-way classification method based on semi-supervised shadowed sets is validated to be superior to produce low-risk classification results on partial labeled data.

Our future works include the following issues. First, the semi-supervised neighborhoods are initialized based on the semi-supervised clustering methods, which require to predefine the number of clusters. We should consider more flexible methods to initialize the neighborhoods based on data distributions. Second, we will design a fast optimization algorithm to speed up solving the objective function of semi-supervised shadowed sets. Finally, the Euclidean distance we adopt for constructing shadowed neighborhoods and classifying uncertain data may be not effective for high-dimensional data. We expect to explore distance metrics to implement a semi-supervised three-way classification method for high-dimensional partial labeled data.

CRediT authorship contribution statement

X.D. Yue: Conceptualization, Methodology. **Q. Qian:** Visualization. **D.Q. Miao:** Writing - review & editing. **C. Gao:** Validation.

Declaration of Competing Interest

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

Acknowledgement

This work was supported by National Natural Science Foundation of China (Serial Nos. 61976134, 61991410, 61991415), National Key Research and Development Program (Serial No. 2018YFB0704400) and Natural Science Foundation of Shanghai (Serial No. 21ZR1423900).

References

[1] W. Pedrycz, Granular computing for data analytics: a manifesto of human-centric computing, IEEE/CAA Journal of Automatica Sinica 5 (6) (2018) 1025–1034.

[2] Y. Yao, L. Zhao, A measurement theory view on the granularity of partitions, Information Sciences 213 (2012) 1–13.

[3] W. Pedrycz, Shadowed sets: representing and processing fuzzy sets, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 28 (1) (1998) 103–109.

[4] W. Pedrycz, From fuzzy sets to shadowed sets: interpretation and computing, International Journal of Intelligent Systems 24 (1) (2009) 48–61.

[5] Y. Yao, S. Wang, X. Deng, Constructing shadowed sets and three-way approximations of fuzzy sets, Information Sciences 412 (2017) 132–153.

[6] C. Li, J. Yi, H. Wang, G. Zhang, J. Li, Interval data driven construction of shadowed sets with application to linguistic word modelling, Information Sciences 507 (2020) 503–521.

[7] Q. Zhang, Y. Chen, J. Yang, G. Wang, Fuzzy entropy: A more comprehensible perspective for interval shadowed sets of fuzzy sets, IEEE Transactions on Fuzzy Systems 28 (11) (2019) 3008–3022.

[8] C. Wu, Q. Zhang, F. Zhao, Y. Cheng, G. Wang, Three-way recommendation model based on shadowed set with uncertainty invariance, International Journal of Approximate Reasoning 135 (2021) 53–70.

[9] S. He, X. Pan, Y. Wang, A shadowed set-based todim method and its application to large-scale group decision making, Information Sciences 544 (2021) 135–154.

[10] H. Zhang, T. Zhang, W. Pedrycz, C. Zhao, D. Miao, Improved adaptive image retrieval with the use of shadowed sets, Pattern Recognition 90 (2019) 390–403.

[11] H. Zheng, Y. Chen, X. Yue, Deep pancreas segmentation with uncertain regions of shadowed sets, Magnetic Resonance Imaging 68 (2020) 45–52.

[12] Y. Yao, Three-way decisions with probabilistic rough sets, Information sciences 180 (3) (2010) 341–353.

[13] Y. Yao, Three-way decision and granular computing, International Journal of Approximate Reasoning 103 (2018) 107–123.

[14] H. Tahayori, A. Sadeghian, W. Pedrycz, Induction of shadowed sets based on the gradual grade of fuzziness, IEEE Transactions on Fuzzy Systems 21 (5) (2013) 937–949.

[15] J. Zhou, D. Miao, C. Gao, Z. Lai, X. Yue, Constrained three-way approximations of fuzzy sets: From the perspective of minimal distance, Information Sciences 502 (2019) 247–267.

[16] Y. Zhang, J. Yao, Game theoretic approach to shadowed sets: a three-way tradeoff perspective, Information Sciences 507 (2020) 540–552.

[17] A. Campagner, V. Dorigatti, D. Ciucci, Entropy–based shadowed set approximation of intuitionistic fuzzy sets, International Journal of Intelligent Systems 35 (12) (2020) 2117–2139.

[18] Q. Zhang, M. Gao, F. Zhao, G. Wang, Fuzzy-entropy-based game theoretic shadowed sets: A novel game perspective from uncertainty, IEEE Transactions on Fuzzy Systems PP (99) (2020) 1–1.

[19] E. Ruspini, Numerical methods for fuzzy clustering, Information Sciences 2 (3) (1970) 319–350.

[20] I. Gath, A. Geva, Unsupervised optimal fuzzy clustering, IEEE Transactions on Pattern Analysis and Machine Intelligence 11 (7) (1989) 773–780.

[21] P. Lingras, C. West, Interval set clustering of web users with rough k-means, Journal of Intelligent Information Systems 23 (1) (2004) 5–16.

[22] S. Mitra, W. Pedrycz, B. Barman, Shadowed c-means: integrating fuzzy and rough clustering, Pattern Recognition 43 (4) (2010) 1282–1291.

[23] J. Zhou, Z. Lai, D. Miao, C. Gao, X. Yue, Multigranulation rough-fuzzy clustering based on shadowed sets, Information Sciences 507 (2020) 553–573.

[24] X. Yue, J. Zhou, Y. Yao, D. Miao, Shadowed neighborhoods based on fuzzy rough transformation for three-way classification, IEEE Transactions on Fuzzy Systems 28 (5) (2020) 978–991.

[25] J. Zhou, C. Gao, W. Pedrycz, Z. Lai, X. Yue, Constrained shadowed sets and fast optimization algorithm, International Journal of Intelligent Systems 34 (10) (2019) 2655–2675.

[26] T. Miyato, S.-I. Maeda, M. Koyama, S. Ishii, Virtual adversarial training: a regularization method for supervised and semi-supervised learning, IEEE Transactions on Pattern Analysis and Machine Intelligence 41 (8) (2018) 1979–1993.

[27] J.E. Van Engelen, H.H. Hoos, A survey on semi-supervised learning, Machine Learning 109 (2) (2020) 373–440.

[28] A. Blum, T. Mitchell, Combining labeled and unlabeled data with co-training, in: Proceedings of The 11th Annual Conference on Computational Learning Theory, 1998, pp. 92–100.

[29] T. Joachims, et al., Transductive inference for text classification using support vector machines, in: The 16th International Conference on Machine Learning, Vol. 99, 1999, pp. 200–209.

[30] F.G. Cozman, I. Cohen, M.C. Cirelo, et al., Semi-supervised learning of mixture models, in: The 20th International Conference on Machine Learning, Vol. 4, 2003, p. 24.

[31] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General, System 17 (2–3) (1990) 191–209.

[32] J. Yang, Y. Yao, A three-way decision based construction of shadowed sets from atanassov intuitionistic fuzzy sets, Information Sciences 577 (2021) 1–21.

[33] M. Gao, Q. Zhang, F. Zhao, Mean-entropy-based shadowed sets: A novel three-way approximation of fuzzy sets, International Journal of Approximate Reasoning 120 (2020) 102–124.

[34] J. Zhou, W. Pedrycz, D. Miao, Shadowed sets in the characterization of rough-fuzzy clustering, Pattern Recognition 44 (8) (2011) 1738–1749.

[35] H. Li, L. Zhang, B. Huang, X. Zhou, Sequential three-way decision and granulation for cost-sensitive face recognition, Knowledge-Based Systems 91 (2016) 241–251.

[36] L. Zhang, H. Li, X. Zhou, B. Huang, Sequential three-way decision based on multi-granular autoencoder features, Information Sciences 507 (2020) 630–643.

[37] M.K. Afridi, N. Azam, J. Yao, E. Alanazi, A three-way clustering approach for handling missing data using gtrs, International Journal of Approximate Reasoning 98 (2018) 11–24.

[38] H. Yu, C. Zhang, G. Wang, A tree-based incremental overlapping clustering method using the three-way decision theory, Knowledge-Based Systems 91 (2016) 189–203.

[39] X. Jia, W. Li, L. Shang, A multiphase cost-sensitive learning method based on the multiclass three-way decision-theoretic rough set model, Information Sciences 485 (2019) 248–262.

[40] D. Liu, The effectiveness of three-way classification with interpretable perspective, Information Sciences 567 (2021) 237–255.

[41] X. Yue, Y. Chen, D. Miao, H. Fujita, Fuzzy neighborhood covering for three-way classification, Information Sciences 507 (2020) 795–808.

[42] F. Min, S.-M. Zhang, D. Ciucci, M. Wang, Three-way active learning through clustering selection, International Journal of Machine Learning and Cybernetics 11 (5) (2020) 1033–1046.

[43] J. Yao, N. Azam, Web-based medical decision support systems for three-way medical decision making with game-theoretic rough sets, IEEE Transactions on Fuzzy Systems 23 (1) (2015) 3–15.

[44] Y. Chen, X. Yue, H. Fujita, S. Fu, Three-way decision support for diagnosis on focal liver lesions, Knowledge-Based Systems 127 (2017) 85–99.

[45] C. Gao, J. Zhou, D. Miao, J. Wen, X. Yue, Three-way decision with co-training for partially labeled data, Information Sciences 544 (2021) 500–518.

[46] A. Bouchachia, W. Pedrycz, Data clustering with partial supervision, Data Mining and Knowledge Discovery 12 (1) (2006) 47–78.

[47] X. Zhu, Z. Ghahramani, Learning from labels and unlabeled data with label propagation, Tech Report 3175 (2004) (2002) 237–244.

[48] S. de Vries, D. Thierens, A reliable ensemble based approach to semi-supervised learning, Knowledge-Based Systems 215 (2021) 106738.