Information Sciences 507 (2020) 772–794

Contents lists available at ScienceDirect

Information Sciences

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

Three-way confusion matrix for classification: A measure driven view

Jianfeng Xu[a,b,c,1], Yuanjian Zhang[a,b,1], Duoqian Miao[a,b,*]

[a] Department of Computer Science and Technology, Tongji University, Shanghai, 201804, China
[b] Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai, 201804, China
[c] Software College, Nanchang University, Jiangxi, 330047, China

ARTICLE INFO

Article history:
Received 28 October 2017
Revised 18 June 2019
Accepted 30 June 2019
Available online 11 July 2019

Keywords:
Three-way decisions
Confusion matrix
Measure
Classification
Rough set theory

ABSTRACT

Three-way decisions (3WD) is an important methodology in solving problems with uncertainty. A systematic analysis on three-way based uncertainty measures is conducive to the promotion of three-way decisions. Meanwhile, confusion matrix, with multifaceted views, serves as a fundamental role in evaluating classification performance. In this paper, confusion matrix is endowed with semantics of three-way decisions. A collection of measures are thus deduced and summarized into seven measure modes. We further investigate the formulation of three-way regions from a measure driven view. To satisfy the preferences of stakeholder, two different objective functions are formulated, and each of them can include different combinations of measures. To demonstrate the effectiveness, we generate probabilistic three-way decisions for a wealth of datasets. Compared with Gini coefficient based and Shannon entropy based objective functions, our model can deduce more satisfying three-way regions.

1. Introduction

Three-way Decisions (3WD) [35,36], developed from rough set theory [26], is an effective theory for decision with uncertainty. It simulates the routine of man-machine interactions by firstly introducing a third option, which characterizes the vagueness of concept, and resolving it subsequently. Taking classification problem as an example, it allows objects with relatively sufficient information to be quickly committed, whereas uncertain objects with relatively poor information are deferred for further evaluation. Thus, three non-overlapped regions are partitioned and correspond to different action strategies. Three-way Decisions has been prevailing in a wide range of applications, such as conflict analysis [14], sentiment classification [13,30], image annotation [16], attribute reduction [27,41], clustering [38], incremental learning [21,25], stream computing [32] and other domains [4,9,15,18,29,44].

A fundamental issue in applying Three-way Decisions is how to construct three regions with reasonable interpretations. Under the umbrella of three-way decisions, a collection of granular computing models, including Rough Sets [29,34,35], Fuzzy Sets [39,45], Dempster-Shafer Theory [22–24] and Shadowed Set [2,37], have been extended and customized. In particular, trade-off goals among regions are discussed from different views. Representative achievements include: Cost-sensitive

* Corresponding author at: College of Electornics and Information Engineering, 4800 Cao'an Highway, Shanghai, China
E-mail address: dqmiao@tongji.edu.cn (D. Miao).
[1] Authors contributed equally.

https://doi.org/10.1016/j.ins.2019.06.064
0020-0255/

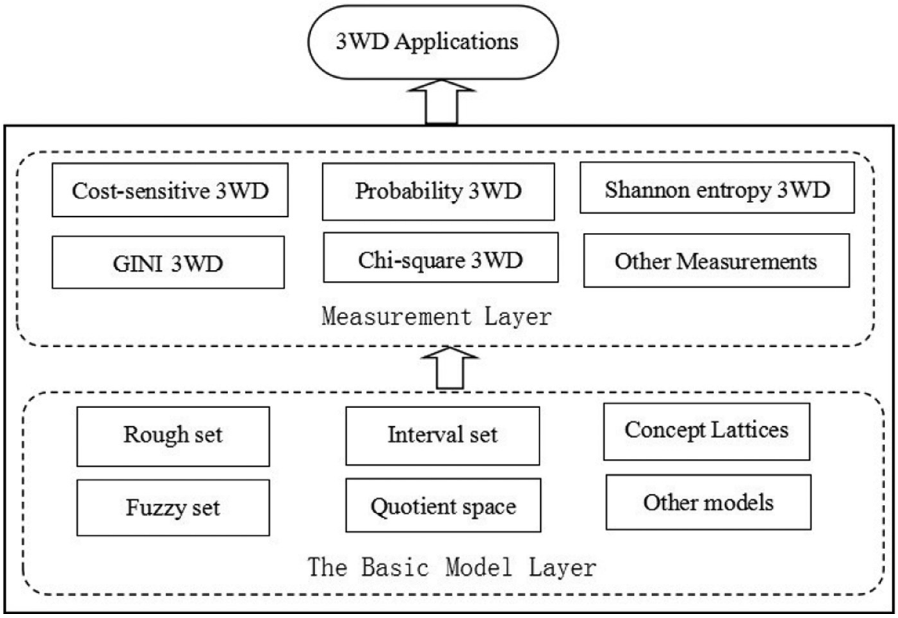


Fig. 1. Hierarchy diagram of the Three-way Decisions evaluation system.

Measure 3WD (Decision-theoretic Rough Set 3WD [3,11,17,29], Fuzzy Set 3WD [5,19], Interval Set 3WD [40], etc.), Conditional Probability Measure 3WD (Probability Rough Set 3WD [45], Game-theoretic Rough Sets 3WD [1,43]), Information Entropy Measure 3WD [6], and GINI Coefficient Measure 3WD [42]. Recent studies suggest that the structure of three-way regions can be approximated sequentially [8], and may be defined in an axiomatic way [28].

The basic evaluation system of 3WD can be viewed as a combination of models and measurements (see Fig. 1). This constitutes the foundation of 3WD reasoning and application. The optimization of thresholds, i.e., obtaining thresholds with numerical solution, is alternative in partitioning three-way regions. In [12], an artificial fish swarm algorithm and an simulated annealing algorithm are proposed to solve the optimal three-way threshold. In [20], a relative value view on risk preference is investigated for determining the region distribution of decision-theoretic rough set. In [33], an optimization mechanism is proposed to deduce an optimal granule representation with balanced test cost and decision cost on rough fuzzy set.

The above studies show that the determination of a three-way regional boundary is influenced by conflicting decision objectives. Therefore, Three-way Decisions threshold solving can be transformed as multi-objective optimization problems [10]. In summary, the main idea of Three-way Decisions is that a certain type of decision evaluation system must be selected firstly, and then a group of appropriate parameters explicitly or implicitly determine the regional boundary of three-way classification. Despite these progress, most of them only extend the formulation from two-way to three-way. The complexity for the classification confusion of a given concept is not quantitatively measured. It means some decision preferences cannot be intuitively formulated.

We attempt to resolve the limitation via confusion matrix. Confusion Matrix [7,31,43] is one of the most classical decision-measure methods in supervised machine learning. It visualizes the degree of algorithm confusion within different classes and is independent of concrete classification algorithm. In view of the significance of Confusion Matrix in the machine learning field, the quantitative analysis of uncertain confusion in performance evaluation can be promoted if it is extended with the semantics of Three-way. For this reason, we deduce a collection of measures in Section 3 and believe that the expansion of measure space can provide more satisfactory performance. Principles for the construction of objective function will be elaborated in Section 4 and rationality of tri-partition will be demonstrated in Section 5.

2. Preliminary: Three-way decisions theory

Three-way Decisions is a trisecting-acting-outcome framework. *Trisecting* means objects are partitioned into three non-overlapped regions. Given a specific measure, a pre-defined evaluation function determines the region affiliation of objects. Objects are classified into the region I, region II and region III for the concept. *Acting* explains the actions we take for the partitioned objects. If we intend to explore the affiliation of objects with regard to a given concept, we can devise three potential options, i.e., acceptance, rejection, and deferment. *Outcome* concerns the measure for the effectiveness of both trisecting and acting. *Tri-partition* and *Acting* operations are reasonable if objects are classified into the ground-truth class as many as possible. The simplicity of the theory is compatible with many granular models, and probabilistic three-way decisions is one of the representatives.

Suppose that the universe of objects U is a finite nonempty set. Let $E \subseteq U \times U$ be an equivalence relation on U, where E is reflexive, symmetric, and transitive, s.t. two objects x and y in U satisfy E (i.e., xEy) if and only if they have the same values across all attributes. The pair $apr = (U, E)$ is called an approximation space. The equivalence relation E induces a partition of U, denoted as $U/E = \{[x] | x \in U\}$, where $[x] = \{y \in U | xEy\}$. For an indescribable target concept $C \subset U$, we can apply conditional probability to measure the probability of object $x \in C$ for any given $[x]$. It is denoted as $Pr(C|[x])$ and can be computed as:

$$Pr(C|[x]) = \frac{Pr([x] \cap C)}{Pr([x])} \tag{1}$$

where $Pr([x])$ denotes the possibility of an arbitrary object x affiliated to $[x]$ given the object set U, and $Pr([x] \cap C)$ denotes the probability of object x affiliated to intersections between an arbitrary equivalence class $[x]$ and C given the object set U. Then we can deduce a probabilistic tri-partition [6,42].

$$\begin{aligned} POS_{(\alpha,\beta)}(C) &= \cup\{[x] \in U/E | Pr(C|[x]) \geq \alpha\}; \\ BND_{(\alpha,\beta)}(C) &= \cup\{[x] \in U/E | \beta < Pr(C|[x]) < \alpha\}; \\ NEG_{(\alpha,\beta)}(C) &= \cup\{[x] \in U/E | Pr(C|[x]) \leq \beta\}. \end{aligned} \tag{2}$$

where $(\alpha, \beta) \in [0, 1] \times [0, 1]$ denotes a pair of thresholds. Regarding this tri-partition as a whole, we have the expressions of a probabilistic three-way decisions $\pi_{(\alpha,\beta)}(C)$ [6,42] for a given concept C:

$$\pi_{(\alpha,\beta)}(C) = \{POS_{(\alpha,\beta)}(C), BND_{(\alpha,\beta)}(C), NEG_{(\alpha,\beta)}(C)\}. \tag{3}$$

Accordingly, we have the acting strategies for classifications of each regions:

- accept $[x]$ with the concept of C if $[x] \subseteq POS_{(\alpha,\beta)}(C)$;
- defer $[x]$ with the concept of C if $[x] \subseteq BND_{(\alpha,\beta)}(C)$;
- reject $[x]$ with the concept of C if $[x] \subseteq NEG_{(\alpha,\beta)}(C)$.

Remark 1. Pawlak rough set and Two-way Decision are two special cases of Three-way Decisions. Pawlak rough set model can be seen as $(\alpha, \beta) = (1.0, 0.0)$, i.e., the decision rules derived from the acceptance or rejection regions are all consistent with 100% accuracy. All inconsistent decision rules are deduced from the deferment region.

In contrast, Two-way decision can be interpreted as $\alpha = \beta$, usually denoted as γ, This model divides all objects into the acceptance and rejection regions. At this point, the conditional probability of decision rules derived from the acceptance region is greater than threshold γ,and the conditional probability of decision rules derived from the rejection region is less than threshold γ.

A variation of thresholds will cause the shape changes of the three-way decision regions, thus simultaneously obtaining relative optimal decisions or balancing the decisions of the three regions is an important research problem in Three-way Decisions theory. In order to evaluate the Three-way Decision measures, a variety of measures such as accuracy, coverage, confidence are proposed for quantitative analysis. In this paper, accuracy rate $ACC(\pi_{(\alpha,\beta)}(C))$ and commitment rate $CMR(\pi_{(\alpha,\beta)}(C))$ are used to measuring the quality of Three-way Decisions.

The accuracy rate is used to evaluate the accuracy degree of decision-making based on the acceptance and rejection regions. It varies in the range of [0, 1]. The formula is as follows [42]:

$$ACC(\pi_{(\alpha,\beta)}(C)) = \frac{|C \cap POS_{(\alpha,\beta)}(C)| + |C^c \cap NEG_{(\alpha,\beta)}(C)|}{|POS_{(\alpha,\beta)}(C)| + |NEG_{(\alpha,\beta)}(C)|}. \tag{4}$$

where $C^c = U - C$.

The commitment rate is used to evaluate the proportion of objects which can be classified with class label via three-way decisions. It varies in the range of [0, 1]. The formula[2] is as follows [42]:

$$CMR(\pi_{(\alpha,\beta)}(C)) = \frac{|POS_{(\alpha,\beta)}(C)| + |NEG_{(\alpha,\beta)}(C)|}{|U|}. \tag{5}$$

The accuracy rate and commitment rate are often contradictory, namely a high accuracy rate is accompanied by a low commitment rate and vice versa. In Pawlak rough set model, for example, the accuracy rate is 1, while the commitment rate is relatively low. Comparatively, a trade-off between accuracy rate and commitment rate is preferred. We believe a measure driven view can facilitate the understanding of three-way decisions and details will be elaborated later.

[2] In [42], it is denoted with the abbreviation *Cov*, we believe *CMR* will be more suitable since *Cov* may be misunderstood as Covariance.

Table 1
An illustration of Three-way Confusion matrix deduced from approximation space $apr = (U, E)$.

	$x \in POS_{(\alpha,\beta)}(C)$	$x \in NEG_{(\alpha,\beta)}(C)$	$x \in BND_{(\alpha,\beta)}(C)$
$x \in C$	$\|POS_{(\alpha,\beta)}(C) \cap C\|$	$\|NEG_{(\alpha,\beta)}(C) \cap C\|$	$\|BND_{(\alpha,\beta)}(C) \cap C\|$
$x \in C^c$	$\|POS_{(\alpha,\beta)}(C) \cap C^c\|$	$\|NEG_{(\alpha,\beta)}(C) \cap C^c\|$	$\|BND_{(\alpha,\beta)}(C) \cap C^c\|$

3. Measure system based on three-way confusion matrix

Confusion Matrix is a visual evaluation tool used in machine learning. The columns of a Confusion Matrix represent the prediction class results, and the rows represent the actual class results. It enumerates all possible cases of a classification problem. Taking binary classification problem as an example, the dimension of a Confusion Matrix is 2×2. A series of algorithm performance measures can be defined using a Confusion Matrix, such as *positive region check rate* and *negative class recall rate*. These measures are generally applicable in all classification algorithms.

When Two-way Decisions is extended to Three-way Decisions, the dimension of the Confusion Matrix of two classification problems is changed from 2×2 to 2×3, That is, the deferment decision class is added to the actual classification result. Table 1 shows a three-way classification Confusion Matrix on determining an arbitrary object x to three-way regions of C in the decision table. The horizontal attributes of the table are the classified object cases. The vertical attributes of the table are the actual object classifications. The intermediate data of the table is the base number of the six sets of objects that are actually positive or negative when they are assigned to the three decision regions.

In Table 1, three-way region is determined by a pair of thresholds (α, β), $x \in U$, $C^c = U - C$, and obviously $Pr(C|[x]) + Pr(C^c|[x]) = 1$, where $[x] \in U/E$, $|POS_{(\alpha,\beta)}(C)| + |NEG_{(\alpha,\beta)}(C)| + |BND_{(\alpha,\beta)}(C)| = |U|$.

3.1. Basic measures of three-way confusion matrix

The meta assessment of classification performance towards Table 1 should be all-encompassed. It is apparent that the vertical perspective focuses on the distribution of the positive against negative for each decision region, whereas the horizontal view focuses on the distribution of the same category in three decision regions. This section will elaborates the measure construction from three perspectives.

The first perspective describes the confusion degrees for positive and negatives classes in each decision region. For an arbitrary decision region, it measures the possibility of instances that are actually affiliated to positive/negative class. Relative measures are defined as follows:

Definition 1. Given a pair of thresholds $(\alpha, \beta) \in [0, 1] \times [0, 1]$, classification uncertainty measures on prediction side include Acceptance Region Accuracy Rate ($M_{PT}(\alpha, \beta)$), Acceptance Region Error Rate ($M_{PF}(\alpha, \beta)$), Rejection Region Accuracy Rate ($M_{NF}(\alpha, \beta)$), Rejection Region Error Rate($M_{NT}(\alpha, \beta)$), Deferment Negative Class Rate ($M_{BF}(\alpha, \beta)$), and Deferment Positive Class Rate ($M_{BT}(\alpha, \beta)$).

Where

$$M_{PT}(\alpha, \beta) = \frac{|POS_{(\alpha,\beta)}(C) \cap C|}{|POS_{(\alpha,\beta)}(C)|}; \tag{6}$$

$$M_{PF}(\alpha, \beta) = \frac{|POS_{(\alpha,\beta)}(C) \cap C^c|}{|POS_{(\alpha,\beta)}(C)|}; \tag{7}$$

$$M_{NF}(\alpha, \beta) = \frac{|NEG_{(\alpha,\beta)}(C) \cap C^c|}{|NEG_{(\alpha,\beta)}(C)|}; \tag{8}$$

$$M_{NT}(\alpha, \beta) = \frac{|NEG_{(\alpha,\beta)}(C) \cap C|}{|NEG_{(\alpha,\beta)}(C)|}; \tag{9}$$

$$M_{BF}(\alpha, \beta) = \frac{|BND_{(\alpha,\beta)}(C) \cap C^c|}{|BND_{(\alpha,\beta)}(C)|}; \tag{10}$$

$$M_{BT}(\alpha, \beta) = \frac{|BND_{(\alpha,\beta)}(C) \cap C|}{|BND_{(\alpha,\beta)}(C)|}. \tag{11}$$

The second perspective corresponds to the confusion degrees for positive and negatives classes in binary class. It measures the possibility of instances in positive/negative class is classified to a particular decision region. Relative measures are defined as follows:

Definition 2. Given a pair of thresholds $(\alpha, \beta) \in [0, 1] \times [0, 1]$, classification uncertainty measures on observation side include Positive Class Recall Rate ($M_{TP}(\alpha, \beta)$), Positive Class Error Rate ($M_{TN}(\alpha, \beta)$), Negative Class Error Rate ($M_{FP}(\alpha, \beta)$),

Table 2
Change analysis of Measure Mode 1 when threshold (α, β) changes from (1.0,0.0).

	(1.0,0.0)	$(\alpha\downarrow, \beta)$	$(\alpha, \beta\uparrow)$	$(\alpha\downarrow, \beta\uparrow)$	(γ, γ)
$M_{TP}(\alpha, \beta)$	*min*	↗	=	↗	↗
$M_{PF}(\alpha, \beta)$	0	↗	=	↗	↗
$M_{FP}(\alpha, \beta)$	0	↗	=	↗	↗
$M_{*P}(\alpha, \beta)$	*min*	↗	=	↗	↗

Negative Class Recall Rate($M_{FN}(\alpha, \beta)$) and Positive Class Deferment Rate ($M_{TB}(\alpha, \beta)$), and Negative Class Deferment Rate ($M_{FB}(\alpha, \beta)$).

Where

$$M_{TP}(\alpha,\beta) = \frac{|POS_{(\alpha,\beta)}(C)\cap C|}{|C|}; \tag{12}$$

$$M_{TN}(\alpha,\beta) = \frac{|NEG_{(\alpha,\beta)}(C)\cap C|}{|C|}; \tag{13}$$

$$M_{FP}(\alpha,\beta) = \frac{|POS_{(\alpha,\beta)}(C)\cap C^c|}{|C^c|}; \tag{14}$$

$$M_{FN}(\alpha,\beta) = \frac{|NEG_{(\alpha,\beta)}(C)\cap C^c|}{|C^c|}; \tag{15}$$

$$M_{TB}(\alpha,\beta) = \frac{|BND_{(\alpha,\beta)}(C)\cap C|}{|C|}; \tag{16}$$

$$M_{FB}(\alpha,\beta) = \frac{|BND_{(\alpha,\beta)}(C)\cap C^c|}{|C^c|}. \tag{17}$$

The third perspective concerns the distribution measure for positive, negative and boundary region in terms of universe. It measures the possibility that instances are determined to certain decision region. Relative measures are defined as follows:

Definition 3. Given a pair of thresholds $(\alpha, \beta)\in[0, 1]\times[0, 1]$, classification uncertainty measures on three-way distribution include Acceptance Region Decision Rate ($M_{*P}(\alpha, \beta)$), Rejection Region Decision Rate ($M_{*N}(\alpha, \beta)$), and Deferment Region Decision Rate($M_{*B}(\alpha, \beta)$). Where

$$M_{*P}(\alpha,\beta) = \frac{|POS_{(\alpha,\beta)}(C)|}{|U|}; \tag{18}$$

$$M_{*N}(\alpha,\beta) = \frac{|NEG_{(\alpha,\beta)}(C)|}{|U|}; \tag{19}$$

$$M_{*B}(\alpha,\beta) = \frac{|BND_{(\alpha,\beta)}(C)|}{|U|}. \tag{20}$$

and symbol $*\in\{P, B, N\}$

By the definition of Three-way Confusion Matrix measures, it can be found that the range of each measure is 0 - 100%. And, the relationship between these measures also has the following properties:

Property 1. *Complementation law of Confusion Matrix measures:*

1.1 $M_{PT}(\alpha,\beta) = 1 - M_{PF}(\alpha,\beta)$
1.2 $M_{NF}(\alpha,\beta) = 1 - M_{NT}(\alpha,\beta)$
1.3 $M_{BT}(\alpha,\beta) = 1 - M_{BF}(\alpha,\beta)$

Proof. According to Definition 1

$$M_{PT}(\alpha,\beta) + M_{PF}(\alpha,\beta) = \frac{|POS_{(\alpha,\beta)}(C)\cap C|}{|POS_{(\alpha,\beta)}(C)|} + \frac{|POS_{(\alpha,\beta)}(C)\cap C^c|}{|POS_{(\alpha,\beta)}(C)|} = \frac{|POS_{(\alpha,\beta)}(C)\cap C| + |POS_{(\alpha,\beta)}(C)\cap C^c|}{|POS_{(\alpha,\beta)}(C)|} = \frac{|POS_{(\alpha,\beta)}(C)\cap(C\cup C^c)|}{|POS_{(\alpha,\beta)}(C)|}$$
$$= \frac{|POS_{(\alpha,\beta)}(C)\cap U|}{|POS_{(\alpha,\beta)}(C)|} = 1$$

Thus, Property 1.1 $M_{PT}(\alpha,\beta) = 1 - M_{PF}(\alpha,\beta)$ is proved.

The proofs of Property 1.2 and 1.3 are similar to that of Property 1.1. □

Table 3
Change analysis of Measure Mode 2 when threshold (α, β) changes from (1.0,0.0).

	(1.0,0.0)	$(\alpha\downarrow, \beta)$	$(\alpha, \beta\uparrow)$	$(\alpha\downarrow, \beta\uparrow)$	(γ, γ)
$M_{NT}(\alpha, \beta)$	0	=	↗	↗	↗
$M_{TN}(\alpha, \beta)$	0	=	↗	↗	↗
$M_{FN}(\alpha, \beta)$	*min*	=	↗	↗	↗
$M_{*N}(\alpha, \beta)$	*min*	=	↗	↗	↗

Table 4
Change analysis of Measure Mode 3 when threshold (α, β) changes from (1.0,0.0).

	(1.0,0.0)	$(\alpha\downarrow, \beta)$	$(\alpha, \beta\uparrow)$	$(\alpha\downarrow, \beta\uparrow)$	(γ, γ)
$M_{PT}(\alpha, \beta)$	1	↘	=	↘	↘

Table 5
Change analysis of Measure Mode 4 when threshold (α, β) changes from (1.0,0.0).

	(1.0,0.0)	$(\alpha\downarrow, \beta)$	$(\alpha, \beta\uparrow)$	$(\alpha\downarrow, \beta\uparrow)$	(γ, γ)
$M_{NF}(\alpha, \beta)$	1	=	↘	↘	↘

Property 2. *Composition Law of Confusion Matrix measures:*

2.1 $M_{TP}(\alpha, \beta) + M_{TN}(\alpha, \beta) + M_{TB}(\alpha, \beta) = 1$
2.2 $M_{FP}(\alpha, \beta) + M_{FN}(\alpha, \beta) + M_{FB}(\alpha, \beta) = 1$

Proof. According to Definitions 2

$$M_{TP}(\alpha, \beta) + M_{TN}(\alpha, \beta) + M_{TB}(\alpha, \beta) = \frac{|POS_{(\alpha,\beta)}(C)\cap C| + |NEG_{(\alpha,\beta)}(C)\cap C| + |BND_{(\alpha,\beta)}(C)\cap C|}{|C|}$$

$$= \frac{|(POS_{(\alpha,\beta)}(C)\cap C)\cup(NEG_{(\alpha,\beta)}(C)\cap C)\cup(BND_{(\alpha,\beta)}(C)\cap C)|}{|D|} = \frac{|POS_{(\alpha,\beta)}(C)\cup NEG_{(\alpha,\beta)}(C)\cup BND_{(\alpha,\beta)}(C))\cap C|}{|C|} = \frac{|U\cap C|}{|C|} = 1$$

Thus, Property 2.1 $M_{TP}(\alpha, \beta) + M_{TN}(\alpha, \beta) + M_{TB}(\alpha, \beta) = 1$ is proved.
The proof of Property 2.2 is similar to that of Property 2.1. □

Property 3. *Integration Law of Confusion Matrix measures:*

3.1 $M_{*P}(\alpha, \beta) = \frac{M_{TP}(\alpha,\beta)\cdot|C| + M_{FP}(\alpha,\beta)\cdot|C^c|}{|U|}$
3.2 $M_{*N}(\alpha, \beta) = \frac{M_{TN}(\alpha,\beta)\cdot|C| + M_{FN}(\alpha,\beta)\cdot|C^c|}{|U|}$
3.3 $M_{*B}(\alpha, \beta) = \frac{M_{TB}(\alpha,\beta)\cdot|C| + M_{FB}(\alpha,\beta)\cdot|C^c|}{|U|}$

Proof. According to Definition 2

$$M_{TP}(\alpha, \beta) = \frac{|POS_{(\alpha,\beta)}(C)\cap C|}{|C|} \quad M_{FP}(\alpha, \beta) = \frac{|POS_{(\alpha,\beta)}(C)\cap C^C|}{|C^c|}$$

Which can be summarized as:

$$\frac{M_{TP}(\alpha,\beta)\cdot|C| + M_{FP}(\alpha,\beta)\cdot|C^c|}{|U|} = \frac{\frac{|POS_{(\alpha,\beta)}(C)\cap C|}{|C|}\cdot|C| + \frac{|POS_{(\alpha,\beta)}(C)\cap C^c|}{|C^c|}\cdot|C^c|}{|U|}$$

$$= \frac{|POS_{(\alpha,\beta)}(C)\cap C| + |POS_{(\alpha,\beta)}(C)\cap C^c|}{|U|} = \frac{|POS_{(\alpha,\beta)}(C)\cap C)\cup(POS_{(\alpha,\beta)}(C)\cap C_j^c)|}{|U|} = \frac{|POS_{(\alpha,\beta)}(C)\cap(C\cup C^c)|}{|U|} = \frac{|POS_{(\alpha,\beta)}(C)\cap U|}{U} = \frac{|POS_{(\alpha,\beta)}(C)|}{|U|}$$

And according to Definition 3

$$M_{*P}(\alpha, \beta) = \frac{|POS_{(\alpha,\beta)}(C)|}{|U|}$$

Thus, Property 3.1 $M_{*P}(\alpha, \beta) = \frac{M_{TP}(\alpha,\beta)\cdot|C| + M_{FP}(\alpha,\beta)\cdot|C^c|}{|U|}$ is proved.
The proofs of Properties 3.2 and 3.3 are similar to that of Property 3.1. □

3.2. Measure modes of three-way confusion matrix

According to the definition of Three-way Decisions Confusion Matrix measures, a change of decision region size will cause a change of the measurement system, and also affect the correct and commitment rates.

Based on the change trend between measured values and thresholds, measures defined in Definitions 1–3 can be divided into seven Measure Modes. In Tables 2–8, the first column corresponds to the measures, the second column lists the values of corresponding measure at the initial thresholds, the other columns list the changes of measure when thresholds change from the initial value(1.0,0.0). There are four possible changes of initial thresholds, i.e., α decreases and β remains 0 $(\alpha\downarrow, \beta)$, α remains 1 and β increases $(\alpha, \beta\uparrow)$, α decreases and β increases $(\alpha\downarrow, \beta\uparrow)$ simultaneously, and α equals to β $(\gamma,$

Table 6
Change analysis of Measure Mode 5 when threshold (α, β) changes from (1.0,0.0).

	(1.0,0.0)	$(\alpha\downarrow, \beta)$	$(\alpha, \beta\uparrow)$	$(\alpha\downarrow, \beta\uparrow)$	(γ, γ)
$M_{TB}(\alpha, \beta)$	*max*	↘	↘	↘	0
$M_{FB}(\alpha, \beta)$	*max*	↘	↘	↘	0
$M_{*B}(\alpha, \beta)$	*max*	↘	↘	↘	0

Table 7
Change analysis of Measure Mode 6 and Measure Mode 7 when threshold (α, β) changes from (1.0,0.0).

	(1.0,0.0)	$(\alpha\downarrow, \beta)$	$(\alpha, \beta\uparrow)$	$(\alpha\downarrow, \beta\uparrow)$	(γ, γ)
$M_{BT}(\alpha, \beta)$	$M_{BT}(\alpha, \beta)$	↘	↗	?	×
$M_{BF}(\alpha, \beta)$	$M_{BF}(\alpha, \beta)$	↗	↘	?	×

Table 8
The thresholds table of getting extreme values in different modes when $0.0 \le \beta \le \alpha \le 1.0$.

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6	Mode 7
Max	(0.0,0.0)	(1.0,1.0)	(1.0,*)	(*,0.0)	(1.0,0.0)	(1.0,1.0)	(0.0,0.0)
Min	(1.0,*)	(*,0.0)	(0.0,0.0)	(1.0,1.0)	(γ, γ)	(0.0,0.0)	(1.0,1.0)

γ). Corresponding measure changes are denoted with symbol "↘" if decreases, "↗" if increases, " = " if unchanged, "?" if uncertain, " × " if value is meaningless, *min* if minimum and *max* if maximum.

Definition 4. For classification uncertainty measures deduced from three-way confusion matrix, Acceptance Region Error Rate ($M_{PF}(\alpha, \beta)$), Negative Class Error Rate ($M_{FP}(\alpha, \beta)$), Positive Class Recall Rate ($M_{TP}(\alpha, \beta)$) and Acceptance Region Decision Rate ($M_{*P}(\alpha, \beta)$) are defined as Measure Mode 1.

We examine the influences of measures for Measure Mode 1 caused by changes of thresholds (α, β).

- Case 1: $(\alpha, \beta) = (1.0, 0.0)$. The acceptance region only contains objects belonging to C, and the negative region only contains objects belonging to C^c. For this reason, we have $POS_{(\alpha,\beta)}(C) = POS_{(\alpha,\beta)}(C) \cap C$. According to Definitions 2 and 3, both Positive Class Recall Rate ($M_{TP}(\alpha, \beta)$) and Acceptance Region Decision Rate($M_{*P}(\alpha, \beta)$) are thus to be the minimal value in the domain w.r.t. α and β. Since $C \cap C^c = \emptyset$, we have $M_{PF}(\alpha, \beta) = 0$ and $M_{FP}(\alpha, \beta) = 0$.
- Case 2: $(\alpha\downarrow, \beta)$. The values for Positive Class Recall Rate ($M_{TP}(\alpha, \beta)$) and Acceptance Region Decision Rate($M_{*P}(\alpha, \beta)$) are increased as more objects with concept C, as compared to objects with concept C^c, are transferred from boundary region of C to positive region of C. The values for Acceptance Region Error Rate ($M_{PF}(\alpha, \beta)$) and Negative Class Error Rate($M_{FP}(\alpha, \beta)$) are increased as some objects with concept C^c are transferred from boundary region of C to positive region of C.
- Case 3: $(\alpha, \beta\uparrow)$. All included measures keep unchanged as the number of objects with/without concept C in positive region of C is unchanged.
- Case 4: $(\alpha\downarrow, \beta\uparrow)$. It can be regarded as a compound operation of Case 2 and Case 3, and trends of measures are increased since both of them are non-decreasing.
- Case 5: $(\alpha, \beta) = (\gamma, \gamma)$. It can be regarded as a compound operation of Case 2 and Case 3, and the final measures are larger since measure trend of both cases expand the positive region.

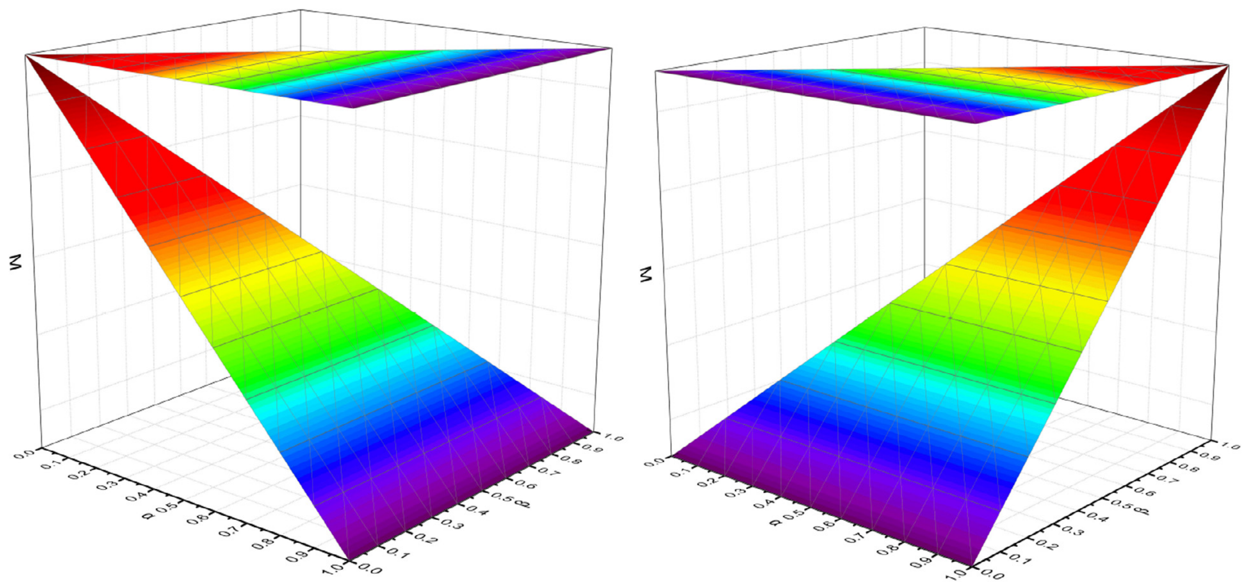
Fig. 2.1 shows the relationship between the thresholds and the measure values in Measure Mode 1.

It can be seen from Fig. 2.1 that, given $0.0 \le \beta \le \alpha \le 1.0$,the measure value of Mode 1 is minimum if $(\alpha, \beta) = (1.0, *)$, and maximum if $(\alpha, \beta) = (0.0, 0.0)$, where * is a wildcard to match any numerical value in the function value range, and the semantic of $(\alpha, \beta) = (1.0, *)$ is $\alpha = 1.0$ and $\beta \in [0, 1]$. The wildcard * in all optimal solutions of thresholds has the same semantics through this paper.

As the assumption $(0.0 \le \beta \le 0.5) \wedge (0.5 \le \alpha \le 1.0)$ is frequently considered in actual classification application, we also discuss the law of variations. From Fig. 2.1, given $(0.0 \le \beta \le 0.5) \wedge (0.5 \le \alpha \le 1.0)$, the measure value of Mode 1 is minimum if $(\alpha, \beta) = (1.0, *)$, and maximum if $(\alpha, \beta) = (0.5, *)$.

Definition 5. For classification uncertainty measures deduced from three-way confusion matrix, Rejection Region Error Rate ($M_{NT}(\alpha, \beta)$), Positive Class Error Rate ($M_{TN}(\alpha, \beta)$), Negative Class Recall Rate ($M_{TP}(\alpha, \beta)$) and Rejection Region Decision Rate ($M_{*N}(\alpha, \beta)$) are defined as Measure Mode 2.

We examine the influences of measures for Measure Mode 2 caused by changes of threshold. (α, β).



2.1 Measures of Mode 1 2.2 Measures of Mode 2

Fig. 2. Comparison of measures of Modes 1 and 2.

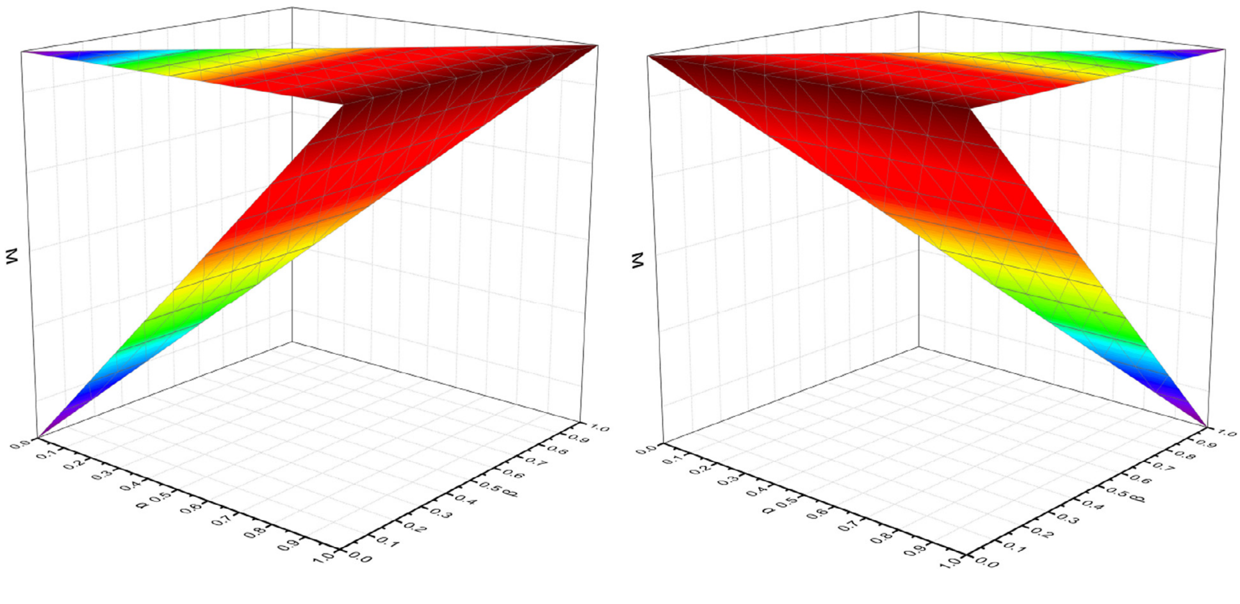
- Case 1: $(\alpha, \beta) = (1.0, 0.0)$. The acceptance region only contains objects belonging to C, and the negative region only contains objects belonging to C^c. For this reason, we have $NEG_{(\alpha,\beta)}(C^c) = NEG_{(\alpha,\beta)}(C^c) \cap C^c$. According to Definitions 2 and 3, both Negative Class Recall Rate ($M_{FN}(\alpha,\ \beta)$) and Rejection Region Decision Rate ($M_{*N}(\alpha,\ \beta)$) are thus to be the minimal value in the domain w.r.t. α and β. Since $C \cap C^c = \emptyset$, we have $M_{NT}(\alpha, \beta) = 0$ and $M_{TN}(\alpha, \beta) = 0$.
- Case 2: $(\alpha\downarrow,\ \beta)$. All included measures keep unchanged as the number of objects with/without concept C in negative region of C is unchanged.
- Case 3: $(\alpha,\ \beta\uparrow)$. The values of Negative Class Recall Rate ($M_{FN}(\alpha,\ \beta)$) and Rejection Region Decision Rate ($M_{*N}(\alpha,\ \beta)$) are increased as more objects with concept C^c, as compared to objects with concept C are transferred from boundary region to negative region. The values of Rejection Region Error Rate ($M_{NT}(\alpha,\ \beta)$) and Positive Class Error Rate ($M_{TN}(\alpha,\ \beta)$), are increased as some objects with concept C are transferred from boundary region of C to negative region of C.
- Case 4: $(\alpha\downarrow,\ \beta\uparrow)$. It can be regarded as a compound operation of Case 2 and Case 3, and trends of measures are increased since variations of measures in both cases are increased.
- Case 5: $(\alpha, \beta) = (\gamma, \gamma)$. It can be regarded as a special case of Case 4, and trends of measures are increased since variations of measures in both cases are increased.

Fig. 2.2 shows the relationship between thresholds and the measure value of Measure Mode 2. It can be seen from Fig. 2.2 that, given $0.0 \le \beta \le \alpha \le 1.0$, the measure value of Mode 2 is minimum if $(\alpha, \beta) = (*, 0.0)$, and maximum if $(\alpha, \beta) = (1.0, 1.0)$. Given $(0.0 \le \beta \le 0.5) \wedge (0.5 \le \alpha \le 1.0)$, the measure value of Mode 2 is minimum if $(\alpha, \beta) = (*, 0.0)$, and maximum if $(\alpha, \beta) = (*, 0.5)$.

Definition 6. For classification uncertainty measures deduced from three-way confusion matrix, Acceptance Region Accuracy Rate ($M_{PT}(\alpha,\ \beta)$) is defined as Measure Mode 3.

We examine the influences of measures for Measure Mode 3 caused by changes of thresholds $(\alpha,\ \beta)$.

- Case 1: $(\alpha, \beta) = (1.0, 0.0)$. The acceptance region only contains objects belonging to C, and the negative region only contains objects belonging to C^c. For this reason, we have $POS_{(\alpha,\beta)}(C) = POS_{(\alpha,\beta)}(C) \cap C$. According to Definition 1, Acceptance Region Accuracy Rate ($M_{PT}(\alpha,\ \beta)$) is thus to be 1.
- Case 2: $(\alpha\downarrow,\ \beta)$. Acceptance Region Accuracy Rate ($M_{PT}(\alpha,\ \beta)$) decreases as some objects with concept C^c are transferred from boundary region of C to positive region of C.
- Case 3: $(\alpha,\ \beta\uparrow)$. Acceptance Region Accuracy Rate ($M_{PT}(\alpha,\ \beta)$) keeps unchanged as the number of objects with/without concept C is unchanged.
- Case 4: $(\alpha\downarrow,\ \beta\uparrow)$. It can be regarded as a compound operation of Case 2 and Case 3, and trends of measures are decreased since variations of measures in both cases are non-increasing.



3.1 Measures of Mode 3 3.2 Measures of Mode 4

Fig. 3. Comparison of the measures of Modes 3 and 4.

- Case 5: $(\alpha, \beta) = (\gamma, \gamma)$. It can be regarded as a special case of Case 4. and the trends of measures are decreased since variations of measures in both cases are non-increasing.

Fig. 3.1 shows the relationship between the thresholds and the measure value of Measure Mode 3. It can be seen that, when $0.0 \leq \beta \leq \alpha \leq 1.0$, if $(\alpha, \beta) = (1.0, *)$, the measure value of Mode 3 is maximum, and if $(\alpha, \beta) = (0.0, 0.0)$, it is minimum. Given $(0.0 \leq \beta \leq 0.5) \wedge (0.5 \leq \alpha \leq 1.0)$, the measure value of Mode 3 is maximum if $(\alpha, \beta) = (0.5, *)$, and minimum if $(\alpha, \beta) = (0.0, 0.0)$.

Definition 7. For classification uncertainty measures deduced from three-way confusion matrix, Rejection Region Accuracy Rate ($M_{NF}(\alpha, \beta)$) is defined as Measure Mode 4.

We examine the influences of measures for Measure Mode 4 caused by changes of thresholds (α, β).

- Case 1: $(\alpha, \beta) = (1.0, 0.0)$. The acceptance region only contains objects belonging to C, and the negative region only contains objects belonging to C^c. For this reason, we have $NEG_{(\alpha,\beta)}(C) = NEG_{(\alpha,\beta)}(C) \cap C$. According to Definition 1, Rejection Region Accuracy Rate ($M_{NF}(\alpha, \beta)$) is thus to be 1.
- Case 2: $(\alpha\downarrow, \beta)$. Rejection Region Accuracy Rate ($M_{NF}(\alpha, \beta)$) remains unchanged since the number of objects with/without concept C is unchanged.
- Case 3: $(\alpha, \beta\uparrow)$. Rejection Region Accuracy Rate ($M_{NF}(\alpha, \beta)$) is decreased as some objects with concept C are transferred from boundary region of C to negative region of C.
- Case 4: $(\alpha\downarrow, \beta\uparrow)$. It can be regarded as a compound operation of Case 2 and Case 3, and trends of measures are decreased since variations of measures in both cases are non-increasing.
- Case 5: $(\alpha, \beta) = (\gamma, \gamma)$. It can be regarded as a special case of Case 4. and the trends of measures are decreased since variations of measures in both cases are non-increasing.

Fig. 3.2 shows the relationship between the thresholds and the measure value of Measure Mode 4. It can be seen that, given $0.0 \leq \beta \leq \alpha \leq 1.0$,measure value of Mode 4 is maximum if $(\alpha, \beta) = (*, 0.0)$, and minimum if $(\alpha, \beta) = (1.0, 1.0)$. Given $(0.0 \leq \beta \leq 0.5) \wedge (0.5 \leq \alpha \leq 1.0)$, the measure value of Mode 4 is maximum if $(\alpha, \beta) = (*, 0.0)$, and minimum if $(\alpha, \beta) = (*, 0.5)$.

Definition 8. For classification uncertainty measures deduced from three-way confusion matrix, Positive Class Deferment Rate ($M_{TB}(\alpha, \beta)$), Negative Class Deferment Rate ($M_{NB}(\alpha, \beta)$) and Deferment Region Decision Rate ($M_{*B}(\alpha, \beta)$) are defined as Measure Mode 5.

We examine the influences of measures for Measure Mode 5 caused by changes of thresholds (α, β).

- Case 1: $(\alpha, \beta) = (1.0, 0.0)$. The acceptance region only contains objects belonging to C, and the negative region only contains objects belonging to C^c. For this reason, we have cardinality of $BND_{(\alpha,\beta)}(C)$ reaches the maximal value. Ac-

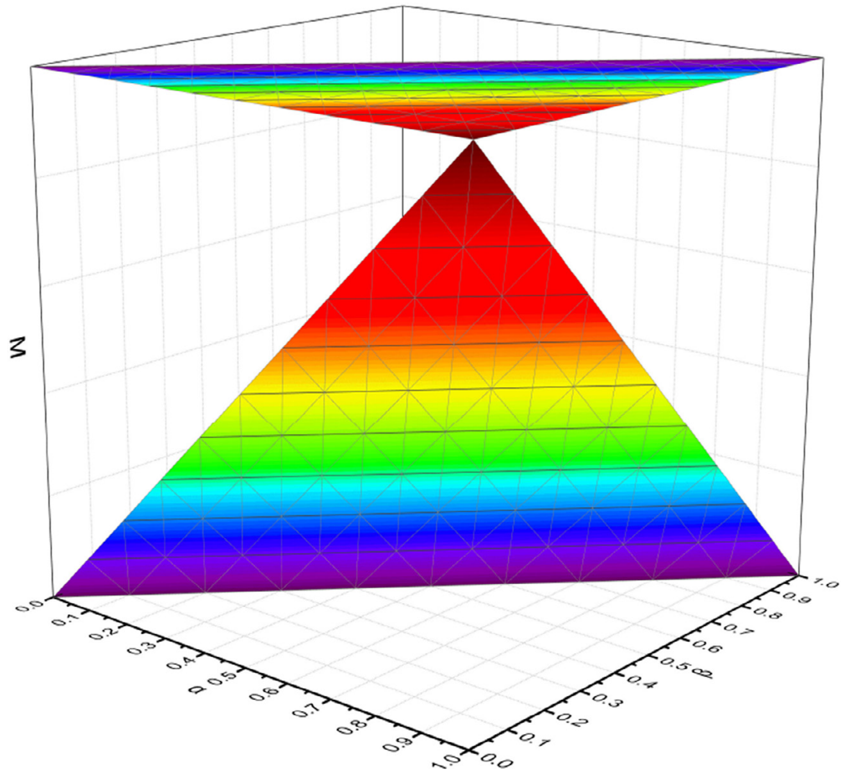


Fig. 4. The value of measures of Mode 5.

cording to Definitions 2 and 3, Positive Class Deferment Rate ($M_{TB}(\alpha,\ \beta)$), Negative Class Deferment Rate ($M_{NB}(\alpha,\ \beta)$), and Deferment Region Decision Rate ($M_{*B}(\alpha,\ \beta)$) are thus to be maximum.

- Case 2: ($\alpha\downarrow,\ \beta$). The values of Positive Class Deferment Rate ($M_{TB}(\alpha,\ \beta)$), Negative Class Deferment Rate ($M_{NB}(\alpha,\ \beta)$), and Deferment Region Decision Rate ($M_{*B}(\alpha,\ \beta)$) are decreased as more objects with concept C, as compared to objects with concept C^c are removed from boundary region to positive region.
- Case 3: ($\alpha,\ \beta\uparrow$). The values of Positive Class Deferment Rate ($M_{TB}(\alpha,\ \beta)$), Negative Class Deferment Rate ($M_{NB}(\alpha,\ \beta)$), and Deferment Region Decision Rate ($M_{*B}(\alpha,\ \beta)$) are decreased as more objects with concept C^c, as compared to objects with concept C, are transferred from boundary region C to negative region C.
- Case 4: ($\alpha\downarrow,\ \beta\uparrow$). It can be regarded as a compound operation of Case 2 and Case 3, and trends of measures are decreased since variations of measures in both cases are decreasing.
- Case 5: $(\alpha, \beta) = (\gamma, \gamma)$. It can be regarded as a special case of Case 4. and the final measures reaches zero since it is not possible for a conditional probability to be larger and smaller than a particular γ meanwhile.

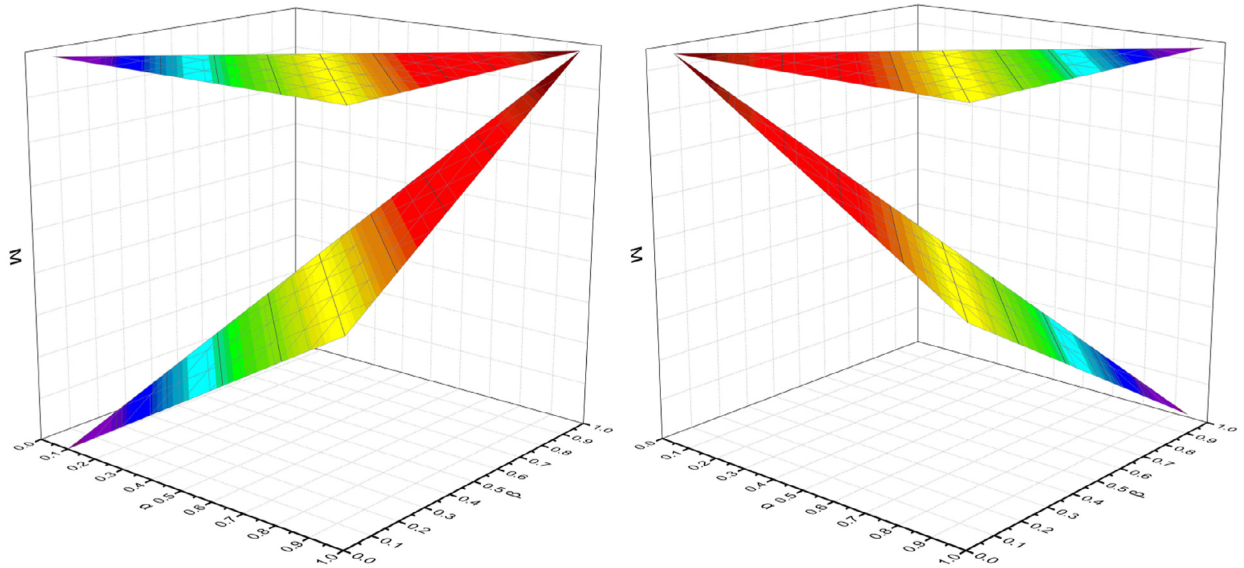
Fig. 4 shows the relationship between the thresholds and the measure value of Measure Mode 5. It can be seen that, given $0.0 \le \beta \le \alpha \le 1.0$, the measure value of Mode 5 is maximum if $(\alpha, \beta) = (1.0, 0.0)$, and minimum if $(\alpha, \beta) = (\gamma, \gamma)$. Given $(0.0 \le \beta \le 0.5) \wedge (0.5 \le \alpha \le 1.0)$,the measure value of Mode 5 is maximum if $(\alpha, \beta) = (1.0, 0.0)$, and minimum if $(\alpha, \beta) = (0.5, 0.5)$.

Definition 9. For classification uncertainty measures deduced from three-way confusion matrix, Deferment Positive Class Rate ($M_{BT}(\alpha,\ \beta)$) is defined as Measure Mode 6.

Definition 10. For classification uncertainty measures deduced from three-way confusion matrix, Deferment Negative Class Rate ($M_{BT}(\alpha,\ \beta)$) is defined as Measure Mode 7.

We examine the influences of measures for Measure Mode 6 and Mode 7 caused by changes of thresholds ($\alpha,\ \beta$).

- Case 1: $(\alpha, \beta) = (1.0, 0.0)$. The acceptance region only contains objects belonging to C, and the negative region only contains objects belonging to C^c. For this reason, we have cardinality of $BND_{(\alpha,\beta)}(C)$ reaches the maximal value. According to Definition 1, Deferment Positive Class Rate ($M_{BT}(\alpha,\ \beta)$) and Deferment Negative Class Rate ($M_{BF}(\alpha,\ \beta)$) may not be either the maximum or minimum. The underlying reason is that the distribution for the conditional probability of equivalence classes is data-dependent.



5.1 Measures of Mode 6

5.2 Measures of Mode 7

Fig. 5. The value of measures of Modes 6 and 7.

- Case 2: $(\alpha\downarrow,\ \beta)$. Equivalence classes are gradually removed from boundary region to positive region if corresponding conditional probabilities w.r.t. concept C are bigger than the decreasing threshold α. This implies that the term $|BND_{(\alpha,\beta)}(C)\cap C|$ is decreased more heavily as compared to $|BND_{(\alpha,\beta)}(C)\cap C^c|$, and thus the Deferment Positive Class Rate ($M_{BT}(\alpha,\ \beta)$) is decreased. Similarly, Deferment Negative Class Rate ($M_{BF}(\alpha,\ \beta)$) is increased.
- Case 3: $(\alpha,\ \beta\uparrow)$. Equivalence classes are gradually removed from boundary region to negative region if corresponding conditional probabilities w.r.t. concept C are smaller than the increasing threshold β. This implies that the term $|BND_{(\alpha,\beta)}(C)\cap C^c|$ is decreased more heavily as compared to $|BND_{(\alpha,\beta)}(C)\cap C|$, and thus the Deferment Negative Class Rate ($M_{BT}(\alpha,\ \beta)$) is decreased. Similarly, Deferment Positive Class Rate ($M_{BF}(\alpha,\ \beta)$) is increased.
- Case 4: $(\alpha\downarrow,\ \beta\uparrow)$. Equivalence classes included in boundary region previously may be removed. It is uncertain for the relation between positive removal and negative removal. Thus, both Deferment Negative Class Rate ($M_{BT}(\alpha,\ \beta)$) and Deferment Positive Class Rate ($M_{BF}(\alpha,\ \beta)$) are uncertain.
- Case 5: $(\alpha,\beta)=(\gamma,\gamma)$. In this case, boundary region becomes empty, and it is illegal to perform division operation.

Fig. 5 shows the relationship between thresholds and the value of measures in measure Mode 6 and 7.

It can be seen from Fig. 5.1 that, given $0.0\le\beta\le\alpha\le1.0$, the measure value of Mode 6 is maximum if threshold $(\alpha,\beta)=(1.0,1.0)$, and minimum if $(\alpha,\beta)=(0.0,0.0)$. Given $(0.0\le\beta\le0.5)\wedge(0.5\le\alpha\le1.0)$, the measure value of Mode 6 is maximum if threshold $(\alpha,\beta)=(1.0,0.5)$, and minimum if $(\alpha,\beta)=(0.5,0.0)$.

It can be seen from Fig. 5.2 that, when $0.0\le\beta\le\alpha\le1.0$, if threshold $(\alpha,\beta)=(1.0,1.0)$, the measure value of Mode 7 is minimum, and if $(\alpha,\beta)=(0.0,0.0)$, its value is maximum. When $(0.0\le\beta\le0.5)\wedge(0.5\le\alpha\le1.0)$, if threshold $(\alpha,\beta)=(1.0,0.5)$, the measure value of Mode 7 is minimum, and if $(\alpha,\beta)=(0.5,0.0)$, its value is maximum.

4. Determining three-way decisions with three-way confusion matrix based measures

An optimal three-way decision is supposed to maximize the expected profits and minimize the expected loss meanwhile for included decision goals. This section intends to explore the formulation of tri-partition by using combinations of measures. The formulation works as heuristic information for searching optimal thresholds (denoted as $(\alpha^*,\ \beta^*)$), which corresponds to an optimal distribution of three-way regions. In this paper, the overall expected objective of Three-way Decisions is expressed by function $T(\alpha,\ \beta)$, the three-way region expected objectives are expressed by $T_p(\alpha,\ \beta)$,$T_B(\alpha,\ \beta)$ and $T_N(\alpha,\ \beta)$, and they are composed of measures proposed by Definitions 1–3, respectively.

Obviously, with a changes of threshold $(\alpha,\ \beta)$, the distribution of objects in the three-way regions will change synchronously, and each measure will also change. However, the dynamic trends of some measures are contradictory, while others may share common dynamic trends. In this paper, two types of Confusion Matrix objective functions are formulated: 1) the expectation objective function of weighting summation method, and 2) the expectation objective function of considering relevant measures simultaneously. This study thus aims to obtain optimal three-way classification by computing objective functions.

4.1. Weighting summation method

The expectation objective function of weighting summation method has many modalities. This paper focuses on maximizing or minimizing the overall objective function. The weight coefficients of measures in the objective function are $\{-1, 0, 1\}$,representing positive attention, non-attention and negative attention, respectively.

The expectation objective function of weighting summation method can be formulated as:

$$\begin{aligned}
&(\alpha^*, \beta^*) = \{(\alpha, \beta)| max/min\, T(\alpha, \beta)\};\\
&where\\
&T(\alpha, \beta) = \epsilon T_P(\alpha, \beta) + \phi T_N(\alpha, \beta) + \psi T_B(\alpha, \beta)\\
&T_p(\alpha, \beta) = w_1 M_{TP}(\alpha, \beta) + w_2 M_{PT}(\alpha, \beta) + w_3 M_{PF}(\alpha, \beta) + w_4 M_{FP}(\alpha, \beta) + w_5 M_{*P}(\alpha, \beta);\\
&T_N(\alpha, \beta) = w_6 M_{FN}(\alpha, \beta) + w_7 M_{NF}(\alpha, \beta) + w_8 M_{TN}(\alpha, \beta) + w_9 M_{NT}(\alpha, \beta) + w_{10} M_{*N}(\alpha, \beta);\\
&T_B(\alpha, \beta) = w_{11} M_{BT}(\alpha, \beta) + w_{12} M_{BF}(\alpha, \beta) + w_{13} M_{TB}(\alpha, \beta) + w_{14} M_{FB}(\alpha, \beta) + w_{15} M_{*B}(\alpha, \beta);\\
&s.t. \epsilon, \phi, \psi \in \{0, 1\}, \quad 0.0 \le \beta \le \alpha \le 1.0 \quad or \quad 0.0 \le \beta \le 0.5 \wedge 0.5 \le \alpha \le 1.0.
\end{aligned} \tag{21}$$

To perform tri-partition, we need a pair of thresholds (α^*, β^*) satisfying $(\alpha^*, \beta^*) = \arg\max_{\alpha,\beta} / \arg\min_{\alpha,\beta} T(\alpha, \beta)$ indicates thresholds (α, β) which either maximizing or minimizing the expectation objective function $T(\alpha, \beta)$. w_i is weight coefficient for all included measure.

The above objective function $T(\alpha, \beta)$ can be composed of linear combinations of different measures, and the choices of each measure and its weights in $T(\alpha, \beta)$ depend on different application requirements. For those complementary measures (e.g. $M_{PT}(\alpha, \beta)$ and $M_{PF}(\alpha, \beta)$, it is forbidden that the corresponding weights are selected as 1 or −1 simultaneously. Although no explicit restrictions on the weight selection for other meta measures exist, it can be empirically nominated by weighting relative misclassification risk. Suppose we intend to develop an identification system for terrorists, the fine-tuning of weight can be referred to the security level. If the system is employed by government, the high-level security signifies the risk of misclassify a terrorist to a citizen is much larger than misclassify a citizen to a terrorist. In this case, decision maker only concerns relative high Positive Class Recall Rate $M_{TP}(\alpha, \beta)$, low Positive Class Error Rate $M_{TN}(\alpha, \beta)$ and a high deferment positive class rate $M_{BT}(\alpha, \beta)$. Consequently, optimum three-way regions are determined by (α^*, β^*) and can be computed as:

$$(\alpha^*, \beta^*) = \{(\alpha, \beta)| max\,(M_{TP}(\alpha, \beta) - M_{TN}(\alpha, \beta) + M_{BT}(\alpha, \beta))\}$$

Consider the same task on railway station. Although the risk of misclassify a terrorist to a citizen is much larger than misclassify a citizen to a terrorist holds, officers should take into account the burden of interrogating large number of passengers. A revision on objective function is necessary, and one feasible solution is to consider low negative class recall rate $M_{FP}(\alpha, \beta)$. Consequently, optimum three-way regions are determined by (α^*, β^*) and can be computed as:

$$(\alpha^*, \beta^*) = \{(\alpha, \beta)| max\,(M_{TP}(\alpha, \beta) - M_{TN}(\alpha, \beta) + M_{BT}(\alpha, \beta) - M_{FP}(\alpha, \beta))\}$$

If the system is served for a restaurant, rejecting customers in outlandish clothes may spread an awful reputation nearby. Basic security with minimum deferment is more desired. Thus, optimum three-way regions are determined by (α^*, β^*) and can be computed as:

$$(\alpha^*, \beta^*) = \{(\alpha, \beta)| max\,(M_{TP}(\alpha, \beta) + M_{BT}(\alpha, \beta) - M_{FP}(\alpha, \beta) - M_{*B}(\alpha, \beta))\}$$

According to Definitions 4–10, different Measure Modes have different dynamic trends, and these changes and their extreme trends must be analyzed. It is beneficial to solve the objective function $T(\alpha, \beta)$ in analyzing measures dynamic change and extreme values.

Theorem 1. *For an expectation objective function max/min* $T(\alpha, \beta)$ *with* $w_i \in \{0, 1\}$, *if weighting summation method is considered for formulation principle and the measures in the objective function* $T(\alpha, \beta)$ *are all from the* j^{th} *Measure Mode* (see Definitions 4–10, $j = 1, 2, \cdots, 7$), *then we have the following conclusions:*

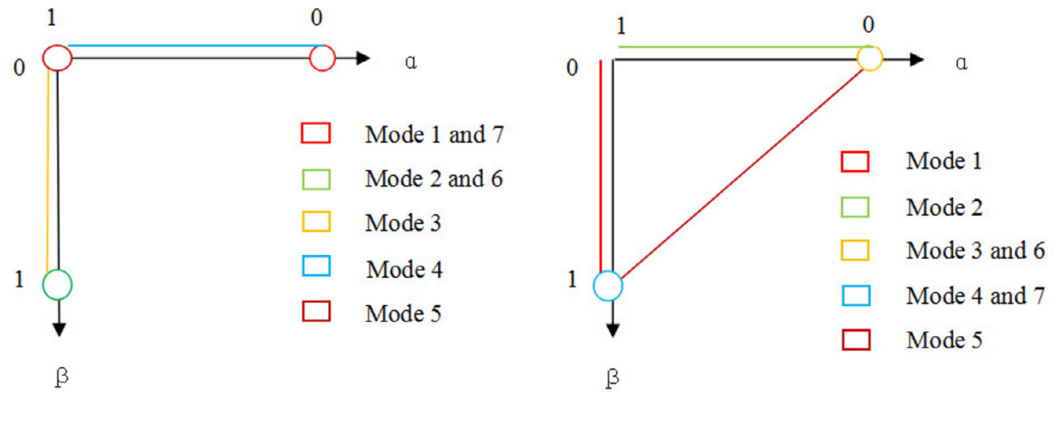
- $(\alpha^*, \beta^*) = \arg\max_{(\alpha,\beta)} T(\alpha, \beta)$, *if for any measures* $M_{**}(\alpha, \beta)$ *included in Measure Mode j,* $M_{**}(\alpha, \beta)$ *reaches its maximum.*
- $(\alpha^*, \beta^*) = \arg\min_{(\alpha,\beta)} T(\alpha, \beta)$, *if for any measures* $M_{**}(\alpha, \beta)$ *included in Measure Mode j,* $M_{**}(\alpha, \beta)$ *reaches its minimum.*

Proof. When these measures in $T(\alpha, \beta)$ belong to the same Measure Mode defined in Definitions 4–9, the dynamic trends of these measures are same. It is also worth mentioning that all measures obtain the minimum or maximum at the same thresholds if all of them are from the same measure mode. Thus, the optimum threshold for the objective function $T(\alpha, \beta)$ is the extreme threshold of the Measure Mode. □

The following Table 8 shows the threshold values for each Measure Mode to take the extreme value when $0.0 \le \beta \le \alpha \le 1.0$. Table 9 shows the threshold values for each Measure Mode in order to take the extreme value when $(0.0 \le \beta \le 0.5) \wedge (0.5 \le \alpha \le 1.0)$, where the maximum threshold represents the threshold when the measure is at maximum,

Table 9
The thresholds table of getting extreme values in different modes when $(0.0 \le \beta \le 0.5) \wedge (0.5 \le \alpha \le 1.0)$.

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6	Mode 7
Max	(0.5,*)	(*,0.5)	(1.0,*)	(*,0.0)	(1.0,0.0)	(1.0,0.5)	(0.5,0.0)
Min	(1.0,*)	(*,0.0)	(0.5,*)	(*,0.5)	(0.5,0.5)	(0.5,0.0)	(1.0,0.5)



6.1 Maximum threshold distribution chart 6.2 Minimum threshold distribution chart

Fig. 6. Threshold distribution chart of extreme values for different Measure Modes when $0.0 \le \beta \le \alpha \le 1.0$.

and the minimum threshold represents the threshold when the measure is at minimum. The objective function satisfying Theorem 1 can quickly yield the expected threshold by searching Tables 8 or 9.

As an example, consider a $T(\alpha, \beta)$ composed of any measures in Measures Mode 1. According to Theorem 1 and Tables 8 and 9, when the threshold range is $0.0 \le \beta \le \alpha \le 1.0$, the threshold of maximal expectation $T(\alpha, \beta)$ is $(\alpha^*, \beta^*) = (0.0, 0.0)$, and the minimum expected threshold is $(\alpha^*, \beta^*) = (1.0, *)$. When the threshold range is $(0.0 \le \beta \le 0.5) \wedge (0.5 \le \alpha \le 1.0)$, the threshold of maximal expectation $T(\alpha, \beta)$ is $(\alpha^*, \beta^*) = (0.5, *)$, and the minimum expected threshold is $(\alpha^*, \beta^*) = (1.0, *)$.

Theorem 2. *For an expectation objective function max/min $T(\alpha, \beta)$ with $w_i \in \{0, 1\}$, if the weighting summation method is considered for formulation principle, and the measures in the objective function $T(\alpha, \beta)$ belong to different Measure Modes (see Definitions 4 to 10), then we have the following conclusions:*

- $(\alpha^*, \beta^*) = \arg\max_{\alpha,\beta} T(\alpha, \beta)$, *if the maximum thresholds of any measure $M_{**}(\alpha, \beta)$ included in $T(\alpha, \beta)$ have non-empty intersection(s) (α^*, β^*);*
- $(\alpha^*, \beta^*) = \arg\min_{\alpha,\beta} T(\alpha, \beta)$, *if the minimum thresholds of any measure $M_{**}(\alpha, \beta)$ included in $T(\alpha, \beta)$ have non-empty intersection(s) (α^*, β^*).*

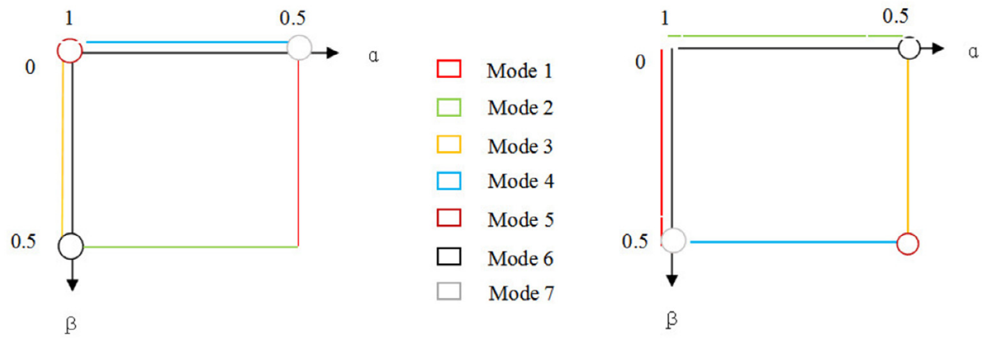
Proof. The proof for first conclusion is given as follows. Let $(\alpha^*, \beta^*) = \bigcap \arg\max_{\alpha,\beta}(M_{**}(\alpha, \beta))$, where $M_{**}(\alpha, \beta)$ is a measure in Definitions 4–10, the sum for the maximum of every measure yields to the maximum of overall target value, which signifies $(\alpha^*, \beta^*) = \arg\max_{\alpha,\beta} T(\alpha, \beta)$. The proof for the second conclusion is similar. □

Fig. 6 is the two-dimensional extreme threshold distribution chart for different Measure Modes, where the α axis scales from 1 to 0, and the β axis scales from 0 to 1, with $0.0 \le \beta \le \alpha \le 1.0$.

Fig. 6.1 is the schematic diagram of the maximum threshold distribution for each Measure Mode. Fig. 6.2 is the schematic diagram of the minimum threshold distribution. As shown in Fig. 6.1, Measure Modes 1, 4 and 6 have threshold intersections, Measure Modes 3, 4 and 5 have threshold intersections, and Measure Modes 3, 2 and 6 have threshold intersections. Fig. 6.2 shows that Measure Modes 2, 3, 5 and 6 have threshold intersections, Measure Modes 1 and 2 have threshold intersections, and Measure Modes 1, 4 and 5 have threshold intersections.

As an example, consider an objective function $T(\alpha, \beta) = M_{TP}(\alpha, \beta) + M_{NF}(\alpha, \beta)$, where $0.0 \le \beta \le \alpha \le 1.0$. In $T(\alpha, \beta)$, $M_{TP}(\alpha, \beta)$ belongs to Measure Mode 1, and $M_{NF}(\alpha, \beta)$ belongs to Measure Mode 4. According to Table 8 or Fig. 6.1, the maximum thresholds for Measure Modes 1 and 4 are (0.0,0.0) and (*, 0.0), because $(0.0, 0.0) \cap (*, 0.0) = (0.0, 0.0)$, and according to Theorem 2, the maximum expected threshold for $T(\alpha, \beta)$ is (0.0,0.0). Similarly, the minimum thresholds for Measure Modes 1 and 4 are (1.0, *) and (1.0,1.0).

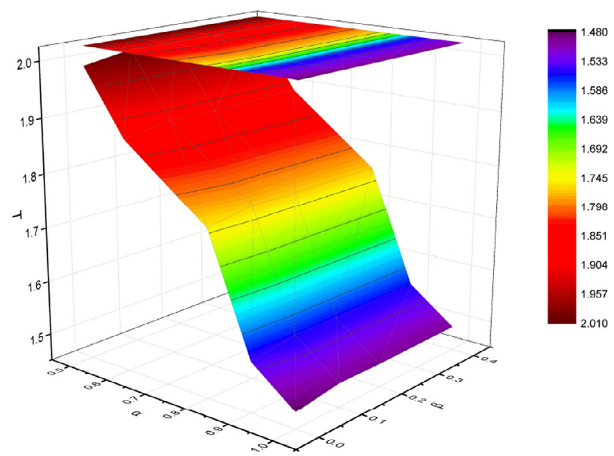
Fig. 7 is the extreme threshold distribution chart for different Measure Modes when $(0.0 \le \beta \le 0.5) \wedge (0.5 \le \alpha \le 1.0)$.Fig. 7.1 is the schematic diagram of the maximum threshold distribution for each Measure Mode. Fig. 7.2 is the schematic diagram



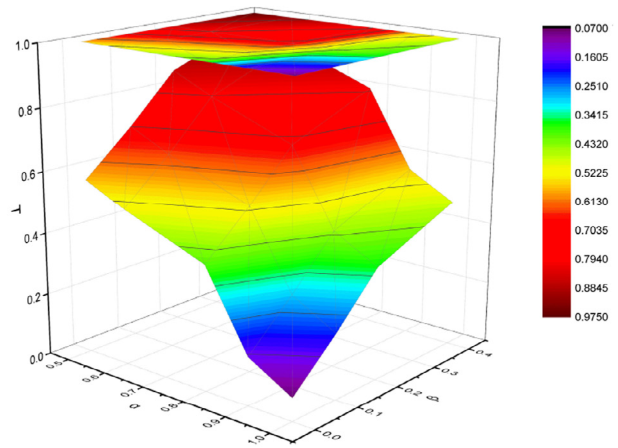
7.1 Maximum threshold distribution chart

7.2 Minimum threshold distribution chart

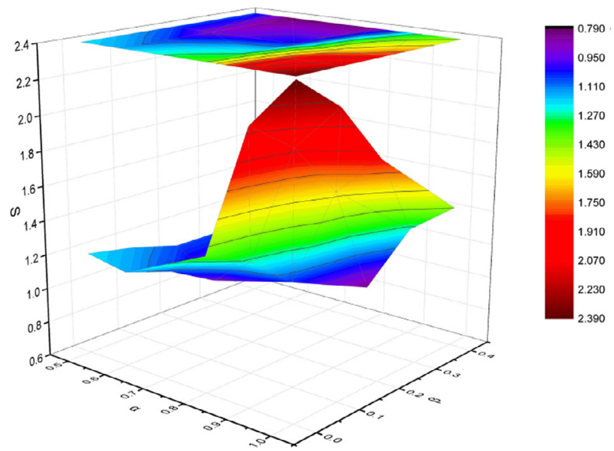
Fig. 7. Threshold distribution chart of extreme values for different Measure Modes when $(0.0 \leq \beta \leq 0.5) \wedge (0.5 \leq \alpha \leq 1.0)$.



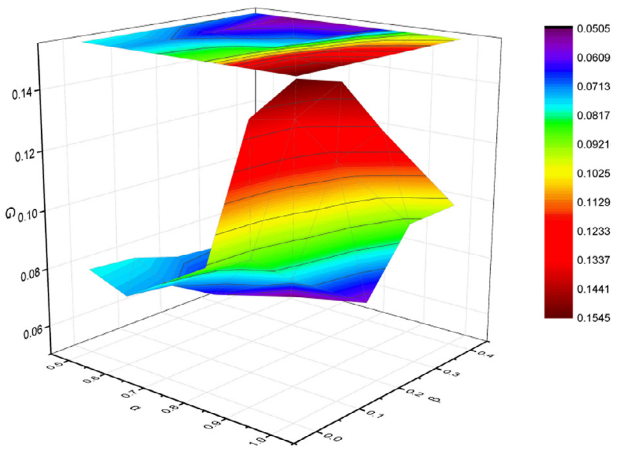
(a) Case 1



(b) Case 2



(c) Entropy



(d) GINI

Fig. 8. Experimental comparison between two three-way decision confusion matrix based functions with information entropy and gini coefficient in dataset 1(*Breast Cancer*), where T,S,G stand for three-way confusion matrix based value, shannon entropy and gini coefficient respectively. All values of T, S, G are determined by (α, β) and are projected for readability. Corresponding values of objective functions are scattered with different colors and displayed in the right-hand side.

of the minimum threshold distribution. In Fig. 7.1, Measure Modes 1 and 4 have threshold intersections, Measure Modes 3, 4 and 5 have threshold intersections, Measure Modes 3, 2 and 6 have threshold intersections and Measure Modes 1 and 2 have threshold intersections.

In Fig. 7.2, Measure Modes 2, 3 and 6 have threshold intersections, Measure Modes 1 and 2 have threshold intersections, Measure Modes 1 and 4 have threshold intersections and Measure Modes 3, 4 and 5 have threshold intersections.

Theorem 3. *For an expectation objective function max/min $T(\alpha,\ \beta)$ with $w_i \in \{-1, 0, 1\}$, if the weighting summation method is considered for formulation principle, and the measures in the objective function $T(\alpha,\ \beta)$ belong to different Measure Modes (see Definitions 4 to 10), then we have the following conclusions:*

- $(\alpha^*, \beta^*) = \arg\max_{\alpha,\beta} T(\alpha, \beta)$, *if all maximal thresholds of measures included in $T(\alpha,\ \beta)$ with weight $= 1$ and all minimum thresholds of measures included in $T(\alpha,\ \beta)$ with weight $= -1$ have non-empty intersection(s) $(\alpha^*,\ \beta^*)$;*
- $(\alpha^*, \beta^*) = \arg\min_{\alpha,\beta} T(\alpha, \beta)$, *if all maximal thresholds of measures included in $T(\alpha,\ \beta)$ with weight $= -1$ and all minimum thresholds of measures included in $T(\alpha,\ \beta)$ with weight $= 1$ have non-empty intersection(s) $(\alpha^*,\ \beta^*)$.*

Proof. We first prove that the theorem holds for maximizing objective function $T(\alpha,\ \beta)$. Let the included measures be partitioned into two sets $M^{+1}(\alpha, \beta)$ and $M^{-1}(\alpha, \beta)$, with the weight of all measures in $M^{+1}(\alpha, \beta)$ be 1 and the weight of all measures in $M^{-1}(\alpha, \beta)$ be -1. Suppose there are two threshold pairs denoted as (α^+, β^+) and (α^-, β^-), and satisfies that (α^+, β^+) and (α^-, β^-) are the intersections that maximizes all measures in $M^{+1}(\alpha, \beta)$ and minimizes all measures in $M^{-1}(\alpha, \beta)$ respectively. Then the sum for the value of measures in $M^{+1}(\alpha, \beta)$ are the maximum and the sum for the values of measures in $M^{-1}(\alpha, \beta)$ are the minimum. If $(\alpha^*, \beta^*) = (\alpha^+, \beta^+) \bigcap (\alpha^-, \beta^-)$, then the sum of values in $M^{+1}(\alpha, \beta)$ minus the sum of values in $M^{-1}(\alpha, \beta)$ reaches maximum. Dually, we can conclude that it holds for minimizing case. □

Many measure modes share extremal thresholds. For example, The maximum thresholds, i.e., $(\alpha^*, \beta^*) = (1.0, 1.0)$ of Measure Modes 2, 3 and 6, and minimum thresholds, i.e., $(\alpha^*, \beta^*) = (1.0, 0.5)$ of Measure Modes 1, 4, and 7 have non-empty intersections.

The above investigations reveal that optimal threshold can be calculated quickly if an expectation objective function based on weighting summation method satisfies Theorems 1, 2 or 3. Otherwise, the optimal threshold must be obtained by computation.

4.2. Considering relevant measures simultaneously

Confusion Matrix measures can be studied simultaneously. The semantics of the expectation objective function of considering relevant measures simultaneously are that these measures of expectation objectives in objective functions should be set separately. This paper mainly consider two objective function patterns.

Pattern 1: The extreme values of each measure are considered respectively

This pattern can be defined as follows:

$$
\begin{aligned}
&(\alpha^*, \beta^*) = \{(\alpha, \beta) | min/max\, T(\alpha, \beta)\}\\
&\textit{where}\\
&T(\alpha, \beta) = T_P(\alpha, \beta) \wedge T_N(\alpha, \beta) \wedge T_B(\alpha, \beta)\\
&T_P(\alpha, \beta) = (w_1 \times (min/max\, M_{TP}(\alpha, \beta))) \wedge (w_2 \times (min/max\, M_{PT}(\alpha, \beta))) \wedge (w_3 \times (min/max\, M_{PF}(\alpha, \beta)))\\
&\qquad \wedge (w_4 \times (min/max\, M_{FP}(\alpha, \beta))) \wedge (w_5 \times (min/max\, M_{*P}(\alpha, \beta)));\\
&T_N(\alpha, \beta) = (w_6 \times (min/max\, M_{FN}(\alpha, \beta))) \wedge (w_7 \times (min/max\, M_{NF}(\alpha, \beta))) \wedge (w_8 \times (min/max\, M_{TN}(\alpha, \beta)))\\
&\qquad \wedge (w_9 \times (min/max\, M_{NT}(\alpha, \beta))) \wedge (w_{10} \times (min/max\, M_{*N}(\alpha, \beta)));\\
&T_B(\alpha, \beta) = (w_{11} \times (min/max\, M_{BT}(\alpha, \beta))) \wedge (w_{12} \times (min/max\, M_{BF}(\alpha, \beta))) \wedge (w_{13} \times (min/max\, M_{TB}(\alpha, \beta)))\\
&\qquad \wedge (w_{14} \times (min/max\, M_{FB}(\alpha, \beta))) \wedge (w_{15} \times (min/max\, M_{*B}(\alpha, \beta)));\\
&s.t.\, 0.0 \le \beta \quad \le \alpha \le 1.0 \quad or \quad 0.0 \le \beta \le 0.5 \wedge 0.5 \le \alpha \le 1.0, \quad w_i \in \{0, 1\} \quad i = 1, 2, \cdots, 15
\end{aligned}
\tag{22}
$$

The description of the above objective function Pattern 1 shows that different measures can be chosen according to their decision demand. It is worth mentioning that the solution may be empty if for all $(\alpha,\ \beta)$, there are always some measures that cannot reach maximum/minimum.

Theorem 4. *For an expectation objective function max/min $T(\alpha,\ \beta)$ with $w_i \in \{0,\ 1\}$, if each measure considered respectively is considered for formulation principle and the measures in the objective function $T(\alpha,\ \beta)$ are all from the j^{th} Measure Mode* (see Definitions 4–10, $j = 1, 2, \cdots, 7$), *then we have the following conclusions:*

- *$(\alpha^*,\ \beta^*)$ is the optimal threshold of $T(\alpha,\ \beta)$, if for any measures $M_{**}(\alpha,\ \beta)$ included in $T(\alpha,\ \beta)$, the maximal values are obtained.*
- *$(\alpha^*,\ \beta^*)$ is the optimal threshold of $T(\alpha,\ \beta)$, if for any measures $M_{**}(\alpha,\ \beta)$ included in $T(\alpha,\ \beta)$, the minimal values are obtained.*

Proof. The proof of Theorem 4 is similar to that of Theorem 1. □

For example, an objective function
$(\alpha^*, \beta^*) = \arg\max_{\alpha,\beta}(M_{TP}(\alpha, \beta)) \bigcap \arg\max_{\alpha,\beta}(M_{PF}(\alpha, \beta)) \bigcap \arg\max_{\alpha,\beta}(M_{FP}(\alpha, \beta))$,
s.t. $(0.0 \le \beta \le 0.5) \wedge (0.5 \le \alpha \le 1.0)$.

Obviously, $M_{TP}(\alpha, \beta)$,$M_{PF}(\alpha, \beta)$ and $M_{FP}(\alpha, \beta)$ belong to Measure Mode 1. As shown in Table 9, the maximum threshold for Measure Mode 1 is (0.5, *). By referring to Theorem 4, the expected threshold for $T(\alpha, \beta)$ is $(\alpha, \beta) = (0.5, *)$.

Theorem 5. *For an expectation objective function* $\max/\min T(\alpha, \beta)$ *with* $w_i \in \{0, 1\}$, *if each measure considered respectively is considered for formulation principle, and the measures in the objective function* $T(\alpha, \beta)$ *belong to different Measure Modes (see Definitions 4 to 10), then we have the following conclusions:*

- (α^*, β^*) *is the optimal thresholds of* $T(\alpha, \beta)$ *if for all measures with* $w_i = 1$, *the values of these measures obtains the desired extreme values (either maximum or minimum)*
- *no solution is available if all measures with* $w_i = 1$, *the intersections of these measures that obtains the desired extreme values (either maximum or minimum) is empty.*

Proof. For the first conclusion, it is similar to that of Theorem 2. For the second conclusion, we cannot find a pair of thresholds (α, β) that satisfies all constraints, thus no solution is found. □

For example, an objective function $T(\alpha, \beta) = max(M_{PT}(\alpha, \beta)) \wedge \min(M_{PF}(\alpha, \beta)) \wedge \min(M_{NF}(\alpha, \beta) \wedge \max(M_{FN}(\alpha, \beta)))$,
s.t. $0.0 \le \beta \le \alpha \le 1.0$.

According to Definitions 4–7, $M_{PT}(\alpha, \beta)$,$M_{PF}(\alpha, \beta)$ and $M_{NF}(\alpha, \beta)$, $M_{FN}(\alpha, \beta)$ belong to Measure Modes 3, 1, 4 and 2, respectively. In Table 8, the maximum threshold for $M_{PT}(\alpha, \beta)$ is (1.0, *), the minimum threshold for $M_{PF}(\alpha, \beta)$ is (1.0, *), the minimum threshold for $M_{NF}(\alpha, \beta)$ is (1.0,1.0), and the maximum threshold for $M_{FN}(\alpha, \beta)$ is (1.0,1.0). Since $(1.0, *) \cap (1.0, 1.0) = (1.0, 1.0)$, according to Theorem 5, the expected threshold for $T(\alpha, \beta)$ is $(\alpha, \beta) = (1.0, 1.0)$.

Pattern 2: Setting the feasible domain for each measure

This pattern can be formulated as:

$$
\begin{aligned}
&(\alpha^*, \beta^*) = \{(\alpha, \beta) | T(\alpha, \beta)\} \\
&where \\
&T(\alpha, \beta) = T_P(\alpha, \beta) \wedge T_N(\alpha, \beta) \wedge T_B(\alpha, \beta) \\
&T_P(\alpha, \beta) = w_1 \times (M_{TP}(\alpha, \beta) \sim t_1) \wedge w_2 \times (M_{PT}(\alpha, \beta) \sim t_2) \wedge w_3 \times (M_{PF}(\alpha, \beta) \sim t_3) \\
&\qquad\qquad \wedge w_4 \times (M_{FP}(\alpha, \beta) \sim t_4) \wedge w_5 \times (M_{*P}(\alpha, \beta) \sim t_5); \\
&T_N(\alpha, \beta) = w_6 \times (M_{FN}(\alpha, \beta) \sim t_6) \wedge w_7 \times (M_{NF}(\alpha, \beta) \sim t_7) \wedge w_8 \times (M_{TN}(\alpha, \beta) \sim t_8) \\
&\qquad\qquad \wedge w_9 \times (M_{NT}(\alpha, \beta) \sim t_9) \wedge w_{10} \times (M_{*N}(\alpha, \beta) \sim t_{10}); \\
&T_B(\alpha, \beta) = w_{11} \times (M_{BT}(\alpha, \beta) \sim t_{11}) \wedge w_{12} \times (M_{BF}(\alpha, \beta) \sim t_{12}) \wedge w_{13} \times (M_{TB}(\alpha, \beta) \sim t_{13}) \\
&\qquad\qquad \wedge w_{14} \times (M_{FB}(\alpha, \beta) \sim t_{14}) \wedge w_{15} \times (M_{*B}(\alpha, \beta) \sim t_{15}); \\
&s.t. 0.0 \le \beta \le \alpha \le 1.0 \, or \, 0.0 \le \beta \le 0.5 \wedge 0.5 \le \alpha \le 1.0, \; w_i \in \{0, 1\}, \; t_i \in [0, 1], \; i = 1, 2, \cdots, 15
\end{aligned}
\tag{23}
$$

Where symbol "$\sim$" represents the relational operators $\le$ or $\ge$, and t_{**} represents the target value range, with $* \in \{P, B, N\}$ if the region is restricted to positive, boundary or negative and can be arbitrary region if not specified. t_i is the lower/upper bound for every considered measure.

The feasible domain of each measure in Pattern 2 is usually specified by the user. The optimal threshold of the objective function Pattern 2 must be obtained by the competitions of various measures. The solution may be solvable or unsolvable, depending on the choice of measures and the setting of the range under different application backgrounds, since the choice of measures and the setting of the range have specific semantics. How to determine the feasible domain is beyond the scope of this paper.

5. Experiment and analysis

This section will illustrate the solution of the expectation function by a simple example first, and evaluate the quality of deduced three-way regions, whose quality are critically compared against alternative tri-partition based on Gini and Shannon Entropy.

5.1. Case study

The optimal expectation thresholds for the Three-way Decisions objective function can be obtained directly when Theorems 1–5 are satisfied. However, a more general objective function cannot satisfy the conditions of Theorems 1–5. For this case, we can search the optimal thresholds by finding all possible pairs of thresholds and store them in a matrix form. Table 10 shows the statistical information to the datasets of with a partition consisting 15 equivalence classes, with X_k represents the *kth* equivalence class. $C \subset U$ denotes a subset of objects with the same label. The conditional probabilities for

Table 10
Data sets information [6] .

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
$Pr(X_i)$	0.0177	0.1285	0.0137	0.1352	0.0580	0.0069	0.0498	0.1070
$Pr(C\|X_i)$	1.0	1.0	1.0	1.0	0.9	0.8	0.8	0.6
	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	
$Pr(X_i)$	0.1155	0.0792	0.0998	0.1299	0.0080	0.0441	0.0067	
$Pr(C\|X_i)$	0.5	0.4	0.4	0.2	0.1	0.0	0.0	

Table 11
The objective function values table of Case 1. Objective function with biggest value is highlighted in boldface.

$\beta\alpha$	1	0.9	0.8	0.6	0.5
0	1.9675	2.0124	2.0492	2.1065	**2.1499**
0.1	1.9587	2.0035	2.0402	2.0979	2.1419
0.2	1.9034	1.9461	1.9808	2.0427	2.0975
0.4	1.8648	1.8983	1.9213	1.9781	1.5719

Table 12
The objective function values of Case 2 under different thresholds. Objective function with smallest value is highlighted in boldface.

$\beta\alpha$	1	0.9	0.8	0.6	0.5
0	0.6128	0.6830	0.7275	0.7216	0.6667
0.1	0.6301	0.7003	0.7448	0.7389	0.6840
0.2	0.8561	0.9263	0.9708	0.9649	0.9100
0.4	1.0169	1.0871	**1.1316**	1.1257	1.0708

decision rules are denoted as $Pr(C|X_k) = \frac{|X_k \cap C|}{|X_k|}$, where $k \in \{1, 2, \cdots, 15\}$. In order to simplify the discussion, the threshold range of the Three-way Decisions is $(0.0 \le \beta \le 0.5) \wedge (0.5 \le \alpha \le 1.0)$.

According to the value information provided in Table 10, and the threshold constraint condition specified in this instance, the threshold value may be $\alpha \in \{1, 0.9, 0.8, 0.6, 0.5\}$, $\beta \in \{0.0, 0.1, 0.2, 0.4\}$.

Case 1. $(\alpha^*, \beta^*) = \{(\alpha, \beta) | max(M_{TP}(\alpha, \beta) + M_{NF}(\alpha, \beta) + M_{BT}(\alpha, \beta))\}$.

Obviously, this expectation objective function of weighting summation method does not satisfy the conditions of Theorems 1–3. According to Definitions 1–3, the values of the included measures ($M_{TP}(\alpha, \beta)$, $M_{NF}(\alpha, \beta)$, $M_{BT}(\alpha, \beta)$) can be calculated at different thresholds, and then weighted into the objective function $T(\alpha, \beta)$. Table 11 shows the values of $T(\alpha, \beta)$ with different thresholds.

As shown in Table 11, the maximum value $T(\alpha, \beta)$ is 2.1499 when $(\alpha, \beta) = (0.5, 0.0)$. Thus, we have $(\alpha^*, \beta^*) = (0.5, 0.0)$.

Case 2. $(\alpha^*, \beta^*) = \{(\alpha, \beta) | min(M_{TP}(\alpha, \beta) - M_{FP}(\alpha, \beta) + M_{FN}(\alpha, \beta) - M_{TN}(\alpha, \beta))\}$.

Obviously, Case 2 does not satisfy Theorems 1–3. According to Definitions 1–3, the values of the included measures ($M_{TP}(\alpha, \beta)$, $M_{FP}(\alpha, \beta)$, $M_{FN}(\alpha, \beta)$, $M_{TN}(\alpha, \beta)$) in $T(\alpha, \beta)$ can be calculated at different thresholds, and then weighted into $T(\alpha, \beta)$.

As shown in Table 12, the minimum $T(\alpha, \beta)$ value is 0.6128 when $(\alpha, \beta) = (1.0, 0.0)$. Thus, we have $(\alpha^*, \beta^*) = (0.8, 0.4)$.

Case 3. $(\alpha^*, \beta^*) = \{(\alpha, \beta) | M_{TP}(\alpha, \beta) \ge 0.9 \wedge M_{PT}(\alpha, \beta) \ge 0.6 \wedge M_{*B}(\alpha, \beta) \ge 0.5\}$.

The $T(\alpha, \beta)$ of Case 3 must evaluate each measure simultaneously. According to Definitions 4, 6 and 8, the measures $M_{TP}(\alpha, \beta)$, $M_{PT}(\alpha, \beta)$, $M_{*B}(\alpha, \beta)$ in Case 3 belong to Modes 1, 3 and 5, respectively. As shown in Table 13, the range of the threshold for $M_{TP}(\alpha, \beta) \ge 0.9$ is $\alpha \in \{1, 0.9, 0.8\}$,$\beta \in \{0.0, 0.1, 0.2, 0.4\}$; the range of the threshold for $M_{PT}(\alpha, \beta) \ge 0.6$ is $\alpha \in \{0.5, 0.6, 0.8\}$ $\beta \in \{0.0, 0.1, 0.2, 0.4\}$; and the range of the threshold for $M_{*B}(\alpha, \beta) \ge 0.5$ is $\alpha \in \{1\}, \beta = \{0.0, 0.1, 0.2\}$ or $\alpha \in \{0.9\}, \beta = \{0.0, 0.1\}$ or $\alpha \in \{0.8\}, \beta = \{0.1, 0.2\}$. Obviously, the threshold that satisfies $T(\alpha, \beta)$ is the intersection of the feasible domain w.r.t. three measures, since:

$\alpha \in \{1, 0.9, 0.8\} \cap \{0.5, 0.6, 0.8\} \cap \{1, 0.9, 0.8\} = \{0.8\}$,

$\beta \in \{0.0, 0.1, 0.2, 0.4\} \cap \{0, 0.1, 0.2, 0.4\} \cap \{0, 0.1, 0.2\} = \{0, 0.1, 0.2\}$

Thus, the optimal thresholds for $T(\alpha, \beta)$ are $(\alpha^*, \beta^*) \in \{(0.8, 0.1), (0.8, 0.2)\}$.

We will examine the distribution of deduced three-way regions via $ACC(\alpha, \beta)$ and $CMR(\alpha, \beta)$ for all cases.

In the first case, threshold $(\alpha, \beta) = (0.5, 0.1)$ satisfies the maximum expected goal. The decision accuracy rate is 82.839%, and the commitment rate is 69.1169%. Threshold $(\alpha, \beta) = (0.5, 0.4)$ satisfies the minimum expected goal, the decision accuracy rate is 78.38% and the commitment rate is 100%. This is equivalent to Two-way Decision with the threshold 0.5. Compared with the minimum expected goal function, the decision commitment rate of the maximum expected goal is decreased to 30.88%, while the accuracy rate is increased by 4.459%.

Table 13
The measure values of Case 3 under different thresholds. Values that satisfy the constraints of every measure are highlighted with underline, and the overall optimal values of objective function are highlighted in boldface.

	$\beta\alpha$	1.0	0.9	0.8	0.6	0.5
$M_{TP}(\alpha,\ \beta)$	0.0	1.0000	0.9836	0.9580	0.8839	0.8137
	0.1	1.0000	0.9836	**0.9580**	0.8839	0.8137
	0.2	1.0000	0.9836	**0.9580**	0.8839	0.8137
	0.4	1.0000	0.9836	0.9580	0.8839	0.8137
$M_{PT}(\alpha,\ \beta)$	0.0	0.4815	0.5667	0.6406	0.7453	0.8395
	0.1	0.4815	0.5667	**0.6406**	0.7453	0.8395
	0.2	0.4815	0.5667	**0.6406**	0.7453	0.8395
	0.4	0.4815	0.5667	0.6406	0.7453	0.8395
$M_{\bullet B}(\alpha,\ \beta)$	0.0	0.6541	0.5961	0.4015	0.2945	0.1790
	0.1	0.6461	0.5881	**0.5394**	0.4324	0.3169
	0.2	0.5162	0.4582	**0.5314**	0.4244	0.3089
	0.4	0.3372	0.2792	0.2225	0.1155	0.0000

Table 14
Comparison of the measures of three-way decisions.

	Shannon entropy	GINI	Confusion matrix
Quantity of measures	1	3	15
Change law of measures	Uncertain	Certain	Certain
The objective function with weighting summation	No	Yes	Yes
The objective function with evaluate each measures simultaneously	No	Yes	Yes

In the second case, threshold $(\alpha, \beta) = (0.8, 0.4)$ satisfies the maximum expected goal. The decision accuracy rate is 85.1318%, and the commitment rate is 77.7578%. Threshold $(\alpha, \beta) = (1, 0)$ satisfies the minimum expected goal, the decision accuracy rate is 100% and the commitment rate is 34.5935%. This is equivalent to a classic Pawlak Three-way Decisions. Compared with Pawlak Three-way Decisions, however, the decision accuracy rate of the maximum expected goal is decreased by 14.86%, while the commitment rate is increased by 43.16%.

In the third case, thresholds $(\alpha, \beta) = (0.8, 0.1)$ satisfies the expected goal function. The decision accuracy rate is 94.81%, and the commitment rate is 46.86%. Compared with Pawlak Three-way Decisions, however, the decision accuracy rate of the maximum expected goal is decreased by 5.19%, while the commitment rate is increased by to 12.27%. $(\alpha, \beta) = (0.8, 0.1)$ satisfies the expected goal function. The decision accuracy rate is 92.66%, and the commitment rate is 59.85%. Compared with Pawlak Three-way Decisions, however, the decision accuracy rate of the maximum expected goal is decreased by 7.34%, while the commitment rate is increased by to 25.26%.

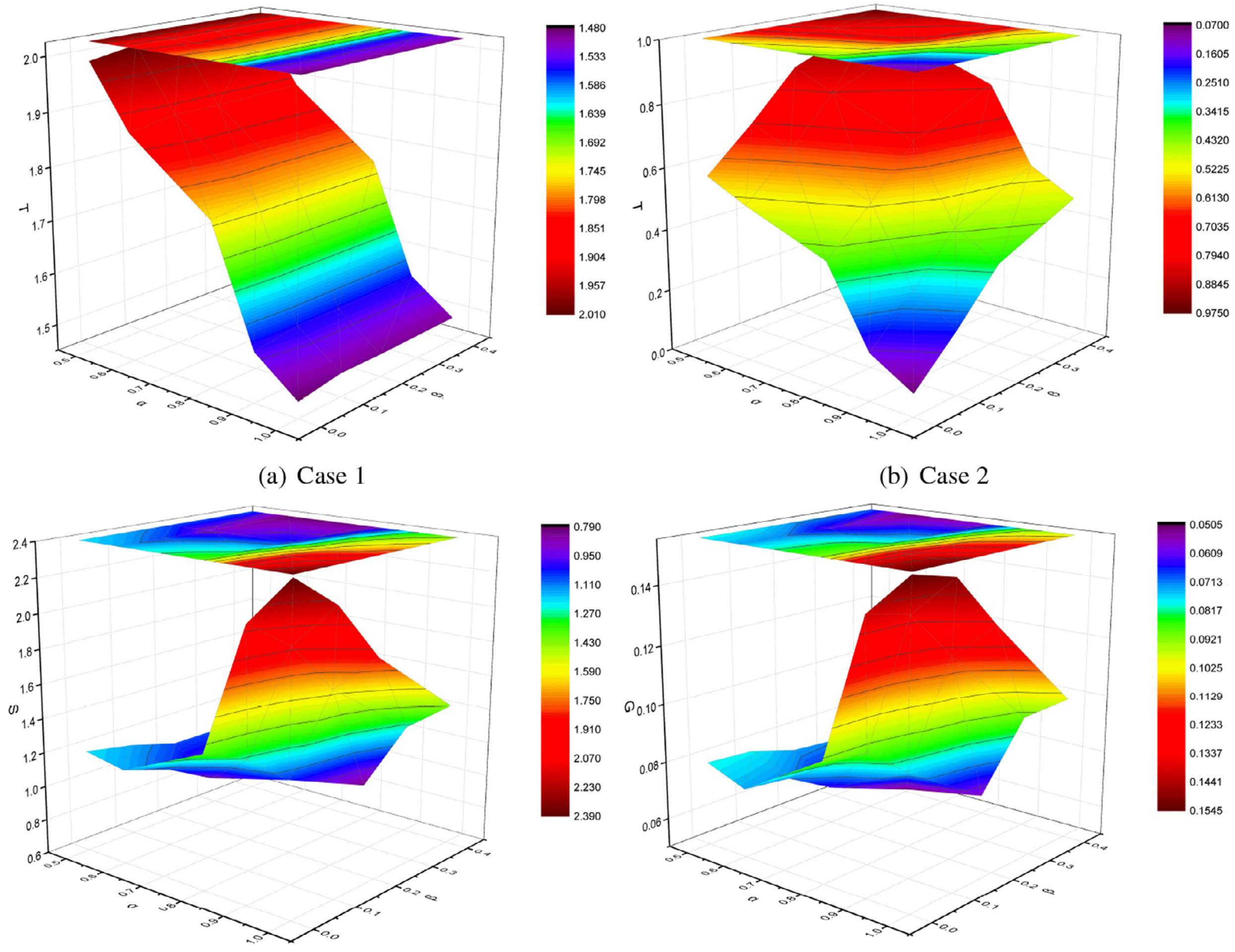
5.2. Comparison of three-way regions from confusion matrix, shannon entropy and GINI

Confusion Matrix, Shannon entropy and GINI coefficient are all used to assess the uncertainty of classification problems. They have often been applied to various fields; in particular, the Confusion Matrix is widely used in machine learning algorithms, the Shannon entropy is used to measure the confusion degree of decisions, while the GINI coefficient is mainly used to measure the impurity of decisions. Shannon entropy and GINI are also important measurement systems in Three-way decision theory. In [6], the overall uncertainty of three regions can be computed as the expected uncertainty, which is referred to as the conditional entropy of π_D, given $\pi_{(\alpha,\beta)}(C)$, specifically, $H(\pi_C|POS_{(\alpha,\cdot)}(C))$, $H(\pi_C|NEG_{(\alpha,\cdot)}(C))$ and $H(\pi_C|POS_{(\alpha,\cdot)}(C))$. In [42], the relative GINI coefficients of the received regions, the rejection regions and the non-commitment rate are expressed as $G_P(\alpha,\ \beta)$, $G_N(\alpha,\ \beta)$ and $G_B(\alpha,\ \beta)$ in three-way decisions.

Table 14 is a comparison of the information of the above three 3DW measurements. As shown in Table 14, Confusion Matrix has the largest number of measures. The measures of Confusion Matrix and GINI coefficient have definite dynamic variation regularity. In terms of objective function form, the Shannon entropy objective function has just one max or min measure expectation for 3WD, and Shannon entropy has only one measure.

The objective function type of the Confusion Matrix is similar to that of the GINI coefficient. However, the measure number of the 3WD Confusion Matrix is more than that of the 3WD GINI coefficient. Therefore, the flexibility and semantic expressive ability of the 3WD Confusion Matrix measures are stronger than those of the 3WD GINI measures.

Table 15 is the dynamic changes comparison table for GINI and partial Confusion Matrix measures. From Table 15, the GINI measure $G_P(\alpha,\ \beta)$ and Confusion Matrix measure $M_{FP}(\alpha,\ \beta)$ that belong to Measure Mode 1 have the consistent dynamic regulation, while the GINI measure $G_N(\alpha,\ \beta)$ and Confusion Matrix measure $M_{NT}(\alpha,\ \beta)$ $M_{TN}(\alpha,\ \beta)$ $M_{FP}(\alpha,\ \beta)$ that belong to Measure Mode 2 have the accordant dynamic regulation. In addition, GINI measure $G_B(\alpha,\ \beta)$ and Confusion Matrix measure $M_{TB}(\alpha,\ \beta)$,$M_{FB}(\alpha,\ \beta)M_{*B}(\alpha,\ \beta)$ that belong to Measure Mode 5 have the identical dynamic regulation.



(a) Case 1 (b) Case 2

(c) Entropy (d) GINI

Fig. 9. Experimental comparison between two three-way decision confusion matrix based functions with information entropy and gini coefficient in dataset 2(*Balance scale*), where T,S,G stands for three-way confusion matrix based value, shannon entropy and gini coefficient respectively. All values of T, S, G are determined by (α, β) and are projected for readability. Corresponding values of objective functions are scattered with different colors and displayed in the right-hand side.

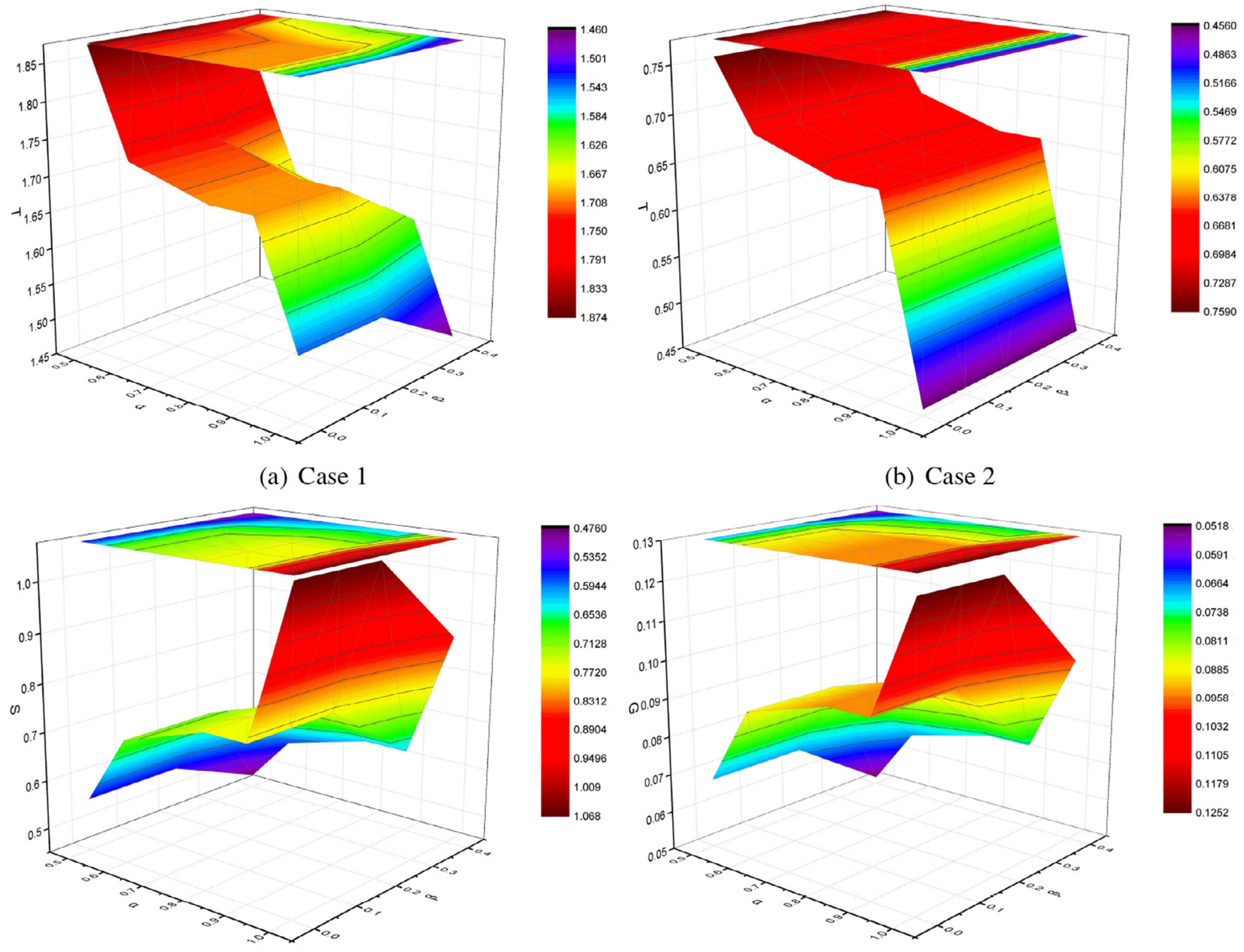
Table 15
The comparison table of GINI and Confusion Matrix measures.

	G_P	M_{PF}, M_{FP}	G_N	M_{NT}, M_{TN}	G_B	M_{TB}, M_{FB}, M_{*B}
$(\alpha, \beta) = (1.0, 0.0)$	0	0	0	0	max	max
$(\alpha\downarrow, \beta)$	$\nearrow$	$\nearrow$	0	0	$\searrow$	$\searrow$
$(\alpha, \beta\uparrow)$	0	0	$\nearrow$	$\nearrow$	$\searrow$	$\searrow$
$(\alpha\downarrow, \beta\uparrow)$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\searrow$	$\searrow$
$(\alpha, \beta) = (\gamma, \gamma)$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	0	0

The above analysis indicates that the dynamic change regulation of GINI measures $G_P(\alpha, \beta)$,$G_N(\alpha, \beta)$ and $G_B(\alpha, \beta)$ accord with matrix Measure Modes 1, 2 and 5, respectively. Therefore, it can be concluded that Theorems 1-5 are also suitable for solving GINI objective functions. Whether the Confusion Matrix measures can be equivalent to the GINI measures, or whether the Confusion Matrix measure can replace the GINI measures remains to be examined.

To examine the prediction performance of three-way confusion matrix based classification, we conduct a group of experiments on four UCI datasets[3]. Detailed characteristics of dataset information are shown in Table 16, where symbol # denotes cardinality.

[3] http://archive.ics.uci.edu/ml/datasets.html



(a) Case 1 (b) Case 2

(c) Entropy (d) GINI

Fig. 10. Experimental comparison between two three-way decision confusion matrix based functions with information entropy and gini coefficient in dataset 3(*Hayes-Roth*), where T,S,G stands for three-way confusion matrix based value, shannon entropy and gini coefficient respectively. All values of T, S, G are determined by (α, β) and are projected for readability. Corresponding values of objective functions are scattered with different colors and displayed in the right-hand side.

Table 16
Characteristics of four real dataset used in experiment.

Name	# entity	# dimensionality	Selected concept (C)	Proportion of positive
Breast cancer	286	9	No-recurrence-events	70.33%
Balance scale	625	4	Left weight	46.08%
Hayes-Roth	160	5	Class 2	42.85%
Car Evaluation	1210	6	Unacceptable	70.02%

For each dataset, we randomly select a subset of objects and investigate the corresponding three-way classification The controlled group include two different objective functions (abbreviated as case 1 and case 2 respectively). The first objective function is formulated as:

$$(\alpha^*, \beta^*) = \{(\alpha, \beta)|(M_{TP}(\alpha, \beta) + M_{NF}(\alpha, \beta) + M_{BT}(\alpha, \beta))\} \tag{24}$$

The second objective function is formulated as:

$$(\alpha^*, \beta^*) = \{(\alpha, \beta)|(M_{TP}(\alpha, \beta) - M_{FP}(\alpha, \beta) + M_{FN}(\alpha, \beta) - M_{TN}(\alpha, \beta))\} \tag{25}$$

For entropy based and gini based function, we use the same settings as discussed in [6,42], as defined in Eqs. (26) and (27):

$$(\alpha^*, \beta^*) = \{(\alpha, \beta)|\big(H(\pi_C|\pi_{\alpha,\beta}(C))\big)\} \tag{26}$$

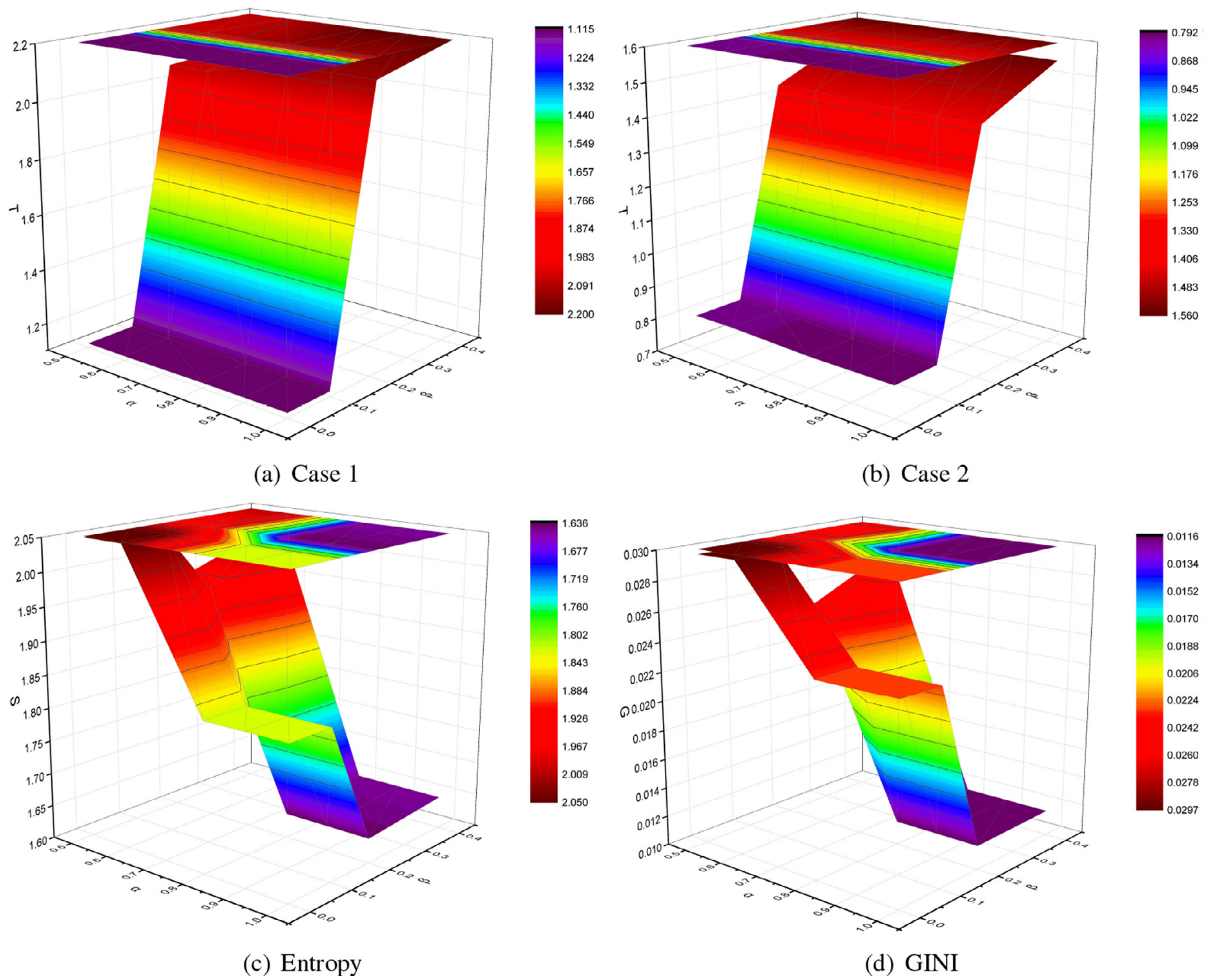


Fig. 11. Experimental comparison between two three-way decision confusion matrix based functions with information entropy and gini coefficient in dataset 4(*Car Evaluation*), where T,S,G stand for three-way confusion matrix based value, shannon entropy and gini coefficient respectively. All values of T, S, G are determined by (α, β) and are projected for readability. Corresponding values of objective functions are scattered with different colors and displayed in the right-hand side.

$$(\alpha^*, \beta^*) = \{(\alpha, \beta) | G_P(\alpha, \beta) + G_N(\alpha, \beta) + G_B(\alpha, \beta)\} \tag{27}$$

Parameters of α, β are increased from 0.5 to 1 and 0 to 0.5 respectively with a step of 0.1. The values of objective functions with the changes of parameter are shown from Fig. 8–11 respectively.

From Figs. 8–11, we can conclude that the distributions of objective function values for entropy and gini are similar, especially for dataset *Breast cancer* and *Balance scale*. This implies that although the definitions of the two metrics are different, yet the value of the optimal threshold may be close. As for proposed objective function (case 1 and case 2), it is interesting to find out that the similarity of two distributions is influenced by dataset. For example, the diversity of case 1 and case 2 is significant for dataset *Breast cancer* and *Balance scale*, whereas is almost identical in terms of dataset *Car Evaluation*. To further investigate the performance of proposed method, we summarize the accuracy rate $ACC(\pi_{(\alpha,\beta)}(C))$ and commitment rate $CMR(\pi_{(\alpha,\beta)}(C))$ of aforementioned methods when the optimal thresholds are obtained in Table 17. Results that achieve the highest $ACC(\pi_{(\alpha,\beta)}(C))$ or $CMR(\pi_{(\alpha,\beta)}(C))$ w.r.t. each dataset are displayed in bold.

It can be inferred from Table 17 that proposed method is likely to achieve an overwhelming commitment rate and a comparable accuracy rate. A possible reason is that for either entropy or gini based solution, balance of three regions is more desired. Therefore, the boundary region is relatively inflated, which in turn increases the possibility of being deferred. Since some ambiguous instances are discarded for decision, the overall accuracy rate is improved. As for proposed method, it is obvious that commitment rate of case 2 is much preferable than case 1, as in case 1 the term $M_{BT}(\alpha, \beta)$ serves as a penalty for true positive being determined. Nevertheless, proposed method is more flexible, as the compared two cases can be replaced by other meta measures. Hence, we claim that presented method embraces the real application more smoothly.

Table 17
Comparison of Accuracy Rate ($ACC(\pi_{(\alpha,\beta)}(C))$) and Commitment Rate ($CMR(\pi_{(\alpha,\beta)}(C))$) in different objective functions given optimal threshold (α, β).

	Objective function of case 1			Objective function of case 2		
	optimal threshold	$ACC(\pi_{(\alpha,\beta)}(C))$	$CMR(\pi_{(\alpha,\beta)}(C))$	optimal threshold	$ACC(\pi_{(\alpha,\beta)}(C))$	$CMR((\pi_{(\alpha,\beta)}(C)))$
Breast cancer	(α, β)=(0.6,0.0)	0.8865	0.8007	(α, β)=(0.5,0.4)	0.9014	**0.9630**
	(α, β)=(0.6,0.1)	0.8865	0.8007			
	(α, β)=(0.6,0.2)	0.8865	0.8007	(α, β)=(0.6,0.4)	0.9014	**0.9630**
Balance scale	(α, β)=(0.5,0.4)	0.7709	**0.8800**	(α, β)=(0.5,0.4)	0.7709	**0.8800**
Hayes-Roth	(α, β)=(0.5,0.0)	**0.8657**	0.5076	(α, β)=(0.5,0.4)	0.8250	**0.6061**
	(α, β)=(0.5,0.1)	**0.8657**	0.5076			
	(α, β)=(0.5,0.2)	**0.8657**	0.5076			
Car evaluation	(α, β)=(0.8,0.4)	0.9160	0.8889	(α, β)=(0.8,0.4)	0.9160	**0.8889**
	(α, β)=(0.9,0.4)	0.9160	0.8889	(α, β)=(0.9,0.4)	0.9160	**0.8889**
	(α, β)=(1.0,0.4)	0.9160	0.8889	(α, β)=(1.0,0.4)	0.9160	**0.8889**
	Objective function of Entropy			Objective function of GINI		
	optimal threshold	$ACC(\pi_{(\alpha,\beta)}(C))$	$CMR(\pi_{(\alpha,\beta)}(C))$	optimal threshold	$ACC((\pi_{(\alpha,\beta)}(C)))$	$CMR((\pi_{(\alpha,\beta)}(C)))$
Breast cancer	(α, β)=(0.9,0.4)	**0.9344**	0.4266	(α, β)=(0.8,0.4)	0.9053	0.6643
Balance scale	(α, β)=(0.6,0.4)	**0.8267**	0.7200	(α, β)=(0.6,0.4)	**0.8267**	0.7200
Hayes-Roth	(α, β)=(0.5,0.4)	0.8250	**0.6061**	(α, β)=(0.5,0.4)	0.8250	**0.6061**
Car evaluation	(α, β)=(0.8,0.2)	**0.9403**	0.8148	(α, β)=(0.8,0.2)	**0.9403**	0.8148
	(α, β)=(0.9,0.2)	**0.9403**	0.8148	(α, β)=(0.9,0.2)	**0.9403**	0.8148
	(α, β)=(1.0,0.2)	**0.9403**	0.8148	(α, β)=(1.0,0.2)	**0.9403**	0.8148

6. Conclusion

In this paper, we have addressed the optimization of Three-way regions systematically from measure view. Firstly, we make an adaptations on confusion matrix with the semantics of three-way decisions. The contained measures are categorized into seven measure modes, and the variations of these measures as thresholds changes are critically analyzed. Secondly, we examine two representative objective functions for determining three-way regions. Compared with the Shannon entropy and GINI measure systems, the Confusion Matrix based measure system and Three-way Confusion Matrix based target function in Three-way Decision provide a more abundant semantic descriptions. These features of the Three-way Decisions Confusion Matrix and the application background of Confusion Matrix itself show that it is more conducive to integration in machine learning. The study of Confusion Matrix measure systems in context of Three-way Decisions reflects the merit of Three-way Decisions in machine learning, and lays the foundation of complicated three-way problem solutions in machine learning. Future work will focus on the concrete application of the Confusion Matrix measure system in Three-way Decisions.

Acknowledgements

Authors would like to thank the anonymous reviewers for their constructive comments and valuable suggestions. This work is supported by the National Natural Science Foundation of China (Nos. 61763031, 61673301, 61573255) National Key R&D Program of China (Grant No. 213), Major Project of Ministry of Public Security (Grant No. 20170004).

References

[1] N. Azam, Y. Zhang, J.T. Yao, Evaluation functions and decision conditions of three-way decisions with game-theoretic rough sets, Eur. J. Oper. Res. 261 (2017) 704–714.
[2] M.J. Cai, Q.G. Li, G.M. Lang, Shadowed sets of dynamic fuzzy sets, Granular Comput. 2 (2) (2017) 85–94.
[3] H.M. Chen, T.R. Li, C. Luo, et al., A decision-theoretic rough set approach for dynamic data mining, IEEE Trans. Fuzzy Syst. 23 (6) (2015) 1958–1970.
[4] Y.F. Chen, X.D. Yue, H. Fujita, S.Y. Fu, Three-way decision support for diagnosis on focal liver lesions, Know.-Based Syst. 127 (2017) 85–99.
[5] X.F. Deng, Y.Y. Yao, Decision-theoretic three-way approximations of fuzzy sets, Inf. Sci. 279 (2014) 702–715.
[6] X.F. Deng, Y.Y. Yao, A multifaceted analysis of probabilistic three-way decisions, Fund. Inf. 132 (2014) 291–313.
[7] X.Y. Deng, Q. Liu, Y. Deng, S. Mahadevan, An improved method to construct basic probability assignment based on the confusion matrix for classification problem, Inf. Sci. (340–341) (2016) 250–261.
[8] Y. Fang, C. Gao, Y.Y. Yao, Granularity-driven sequential three-way decisions: a cost-sensitive approach to classification, Inf. Sci. (2019), doi:10.1016/j.ins.2019.06.003.
[9] L. Hao, F. Tao, J.S. Mi, Three-way decisions for composed set-valued decision tables, J. Intell. Fuzzy Syst. 33 (2) (2017) 937–946.
[10] W.W. Li, X.Y. Jia, L. Wang, B. Zhou, Multi-objective attribute reduction in three-way decision-theoretic rough set model, Int. J. Approx. Reason 105 (2019) 327–341.
[11] F. Jia, P.D. Liu, A novel three-way decision model under multiple-criteria environment, Inf. Sci. 471 (2019) 29–51.
[12] X.Y. Jia, Z.M. Tang, W.H. Liao, et al., On an optimization representation of decision-theoretic rough set model, Int. J. Approx. Reason. 55 (2014) 156–166.
[13] M.T. Khan, N. Azama, S. Khalid, et al., A three-way approach for learning rules in automatic knowledge-based topic models, Int. J. Approx. Reason. 82 (2017) 210–226.
[14] G.M. Lang, D.Q. Miao, H. Fujita, Three-way group conflict analysis based on pythagorean fuzzy set theory, IEEE Trans. Fuzzy Syst. (2019), doi:10.1109/TFUZZ.2019.2908123.
[15] H.X. Li, L.B. Zhang, B. Huang, et al., Sequential three-way decisions and granulation for cost-sensitive face recognition, Knowl. Based Syst. 91 (2016) 241–251.

[16] H.X. Li, L.B. Zhang, X.Z. Zhou, et al., Cost-sensitive sequential three-way decisions modeling using a deep neural network, Int. J. Approx. Reason. 85 (2017) 68–78.
[17] W.T. Li, W.H. Xu, Double-quantitative decision-theoretic rough set, Inf. Sci. 316 (2015) 54–67.
[18] D.C. Liang, W. Pedrycz, D. Liu, et al., Three-way decisions based on decision-theoretic rough sets under linguistic assessment with the aid of group decision making, Appl. Soft Comput. 29 (2015) 256–269.
[19] D.C. Liang, Z.S. Xu, D. Liu, Three-way decisions based on decision-theoretic rough sets with dual hesitant fuzzy information, Inf. Sci. 396 (2017) 127–143.
[20] D.C. Liang, W. Pedrycz, D. Liu, Determining three-way decisions with decision-thoeretic rough sets using a relative value approach, IEEE Trans. Syst. Man Cybern. 47 (8) (2017) 1785–1799.
[21] D. Liu, T.R. Li, J.B. Zhang, Incremental updating approximations in probabilistic rough sets under the variation of attributes, Know. Based Syst. 73 (2015) 81–96.
[22] Z.G. Liu, Q. Pan, J. Dezert, G. Mercier, Hybrid classification system for uncertain data, IEEE Trans. Syst. Man Cybern. Syst. 47 (2017) 2783–2790.
[23] Z.G. Liu, Q. Pan, J. Dezert, A. Martin, Combination of classifiers with optimal weight based on evidential reasoning, IEEE Trans. Fuzzy Syst. 26 (2018) 1217–1230.
[24] Z.G. Liu, Q. Pan, J. Dezert, J.W. Han, Y. He, Classifier fusion with contextual reliability evaluation, IEEE Trans. Cybern. 48 (2018) 1605–1618.
[25] C. Luo, T.R. Li, Y.Y. Yao, Dynamic probabilistic rough sets with incomplete data, Inf. Sci. 4179 (2017) 39–54.
[26] Z. Pawlak, Rough sets, Int. J. Comput. Inf.Sci. 11 (5) (1982) 341–356.
[27] J. Qian, C.Y. Dan, X.D. Yue, et al., Attribute reduction for sequential three-way decisions under dynamic granulation, Int. J. Approx. Reason. 85 (2017) 196–216.
[28] J.S. Qiao, B.Q. Hu, On decision evaluation functions in generalized three-way decision spaces, Inf. Sci. (2018), doi:10.1016/j.ins.2018.07.032.
[29] Y.H. Qian, H. Zhang, Y.L. Sang, et al., Multigranulation decision-theoretic rough sets, Int. J. Approx. Reason. 55 (1) (2014) 225–237.
[30] F.J. Ren, L. Wang, Sentiment analysis of text based on three-way decisions, J. Intell. Fuzzy Syst. 33 (1) (2017) 245–254.
[31] M. Sokolova, G. Lapalme, A systematic analysis of performance measures for classification tasks, Inf. Process. Manag. 45 (4) (2009) 427–437.
[32] J.F. Xu, D.Q. Miao, Y.J. Zhang, et al., A three-way decisions model with probabilistic rough sets for stream computing, Int. J. Approx. Reason. 88 (2017) 1–22.
[33] J. Yang, G.Y. Wang, Q.H. Zhang, et al., Optimal granularity selection based on cost-sensitive sequential three-way decisions with rough fuzzy sets, Know. Based Syst. 163 (2019) 131–144.
[34] Y.Y. Yao, Probabilistic rough set approximations, Int. J. Approx. Reason. 49 (2008) 255–271.
[35] Y.Y. Yao, Three-way decisions with probabilistic rough sets, Inf. Sci. 180 (2010) 341–353.
[36] Y.Y. Yao, The superiority of three-way decisions in probabilistic rough set models, Inf. Sci. 181 (6) (2011) 1080–1096.
[37] Y.Y. Yao, S. Wang, X.F. Deng, Constructing shadowed sets and three-way approximations of fuzzy sets, Inf. Sci. (412–413) (2017) 132–153.
[38] H. Yu, C. Zhang, G.Y. Wang, A tree-based incremental over lapping clustering method using the three-way decisions theory, Know. Based Syst. 91 (2016) 189–203.
[39] L.A. Zadeh, Fuzzy logic = computing with words, IEEE Trans. Fuzzy Syst. 4 (2) (1996) 103–111.
[40] H.Y. Zhang, Y.Y. Shu, M.M. Jian, Ranking interval sets based on inclusion measures and applications to three-way decisions, Know. Based Syst. 91 (2016) 62–70.
[41] X.Y. Zhang, D.Q. Miao, Three-way attribute reducts, Int. J. Approx. Reason. 88 (2017) 401–434.
[42] Y. Zhang, J.T. Yao, GINI Objective functions for three-way classifications, Int. J. Approx. Reason. 81 (2017) 103–114.
[43] Y. Zhang, J.T. Yao, Game theoretic approach to shadowed sets: a three-way tradeoff perspective, Inf. Sci. (2018), doi:10.1016/j.ins.2018.07.058.
[44] Y.J. Zhang, D.Q. Miao, Z.F. Zhang, et al., A three-way selective ensemble model for multi-label classification, Int. J. Approx. Reason. 103 (2018) 394–413.
[45] X.R. Zhao, B.Q. Hu, Fuzzy probabilistic rough sets and their corresponding three-way decisions, Know. Based Syst. 91 (2016) 126–142.