Knowledge-Based Systems 187 (2020) 104804

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

Adaptive weighted generalized multi-granulation interval-valued decision-theoretic rough sets[☆]

Yanting Guo [a], Eric C.C. Tsang [a,*], Weihua Xu [b], Degang Chen [c]

[a] *Faculty of Information Technology, Macau University of Science and Technology, Taipa, Macau, PR China*
[b] *School of Mathematics and Statistics, Southwest University, Chongqing, 400715, PR China*
[c] *Department of Mathematics and Physics, North China Electric Power University, Beijing, 102206, PR China*

ARTICLE INFO

Article history:
Received 13 February 2019
Received in revised form 14 May 2019
Accepted 11 June 2019
Available online 13 June 2019

Keywords:
Interval-valued decision-theoretic rough sets
Weighted generalized multi-granulation
Decision risk

ABSTRACT

With the development of information technology, the sources of information are increasing. How to make good use of information from different sources to make correct decisions is an important problem in multi-source information systems. From the perspective of information granulations, each source can be regarded as a granular structure and the importance of different granulations may be different in multi-source systems. We provide a weighted generalized multi-granulation interval-valued decision-theoretic rough set model (WGM-IVDTRS) for multi-source decision fusion. Firstly, the basic form and important properties of the WGM-IVDTRS model are studied and a granulation weighted method based on the classification accuracy of decision tree learning is proposed from the machine learning point of view. Secondly, three types of the WGM-IVDTRS model are established based on different determination methods of decision risk parameters. Finally, the WGM-IVDTRS models are first compared with multi-granulation decision models and other weighted granulation methods. Moreover, the performances of three WGM-IVDTRS models based on the classification accuracy weighted method are also compared. The experimental comparisons show that the importance, feasibility and effectiveness of the proposed WGM-IVDTRS models, and the third WGM-IVDTRS model performs best when people can accept the range scalability and fault tolerance of intervals.

1. Introduction

Granulation is one of the major concepts of human cognition, which mainly decomposes the whole into parts in the cognitive process [1]. Granulations characterized by different equivalence relations can describe a target concept from different angles. They provide an analytical tool for decision making of multi-source information, multi-intelligence agents, distributed information systems, high-dimensional feature data and so on [2,3]. Nowadays, the information sources of data collected for solving the same problem are increasing, and multi-source data is increasingly common. In this paper, we mainly study the decision-making problem of multi-source decision systems with different attribute sets from the perspective of multi-granulation, where each information source is regarded as a granular structure.

Decision-theoretic rough sets (DTRS) [4,5] based on the Bayesian minimum decision risk provide a cost-effective method for decision-making under single granulation. As a representative decision generalization model of rough sets [2], DTRS has not only made remarkable progress in theory [6–8], but also been widely used in text classification [4], oil exploitation [9], policy decisions [10], web-based medical decision support systems [11], email filtering [12] and so on. When the probability distribution of objects about a target concept in universe is determined, the loss function of DTRS will play a decisive role in decision-making. Some uncertainty factors including the uncertainty of decision-making environment, the urgency of time, the incompleteness of information acquisition and the limitations of the knowledge of decision makers make the loss function more and more inaccurate. Researchers usually express it by using fuzzy forms such as interval number, the trapezoidal fuzzy number and fuzzy number cut sets [13–15]. In particular, some researchers pay attention to the study of DTRS under the interval-valued loss function. Liu et al. [15] first proposed interval-valued decision-theoretic rough sets (IVDTRS). Liang and Liu [16] proposed θ certain ranking method, the degree of possibility ranking method and the optimization method for obtaining three-way decisions in the IVDTRS model. Considering that the minimum decision risk process of IVDTRS is actually a process of sorting intervals, in this paper we first introduce the geometry average interval sorting method [17]

[☆] No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.knosys.2019.06.012.

* Corresponding author.
E-mail address: cctsang@must.edu.mo (E.C.C. Tsang).

https://doi.org/10.1016/j.knosys.2019.06.012
0950-7051/

into the IVDTRS model. Then we provide a specific method to construct decision model under the interval-valued loss function and single granulation environments, which is the geometry average sorting interval-valued decision-theoretic rough set model (IVDTRSG).

Fusion of decisions derived from different granulations is a very important research topic in multi-source decision systems. Without considering the importance of granulations or treating granulations equally, many researchers first proposed multi-granulation decision-theoretic rough set models from optimistic, pessimistic and mean risk preferences. Qian et al. [18] established optimistic, pessimistic and mean multi-granulation decision-theoretic rough set models. Feng et al. [19] presented two variable precision multi-granulation fuzzy decision-theoretic rough sets based on maximal and minimal membership degrees. Yu et al. [20] established three types of double-quantitative decision-theoretic models in multi-granulation approximation spaces. Sun et al. [21] studied multi-granulation fuzzy decision-theoretic rough sets over two universes. Li et al. [22] and Liu et al. [23] studied multi-granulation decision-theoretic rough sets in dominance spaces and multi-covering spaces, respectively. Lin et al. [24] and Li et al. [25] constructed multi-granulation decision methods for multi-source fuzzy information systems and distributed fc-decision information systems, respectively. Zhan et al. [26] and Jiang et al. [27] studied covering based multi-granulation $(\mathcal{I},\mathcal{T})-$fuzzy rough sets and covering based variable precision $(\mathcal{I},\mathcal{T})-$fuzzy rough sets for multi-attribute group decision-making with fuzzy information. In order to enhance the practical application, Xu et al. [28] proposed generalized multi-granulation double-quantitative decision-theoretic rough sets by adopting the majority principle to fuse decisions from different granulations.

The importance of granulations is usually different for decision-making in real life. For example, the evaluations from audiences, singing experts and record companies have different degrees of support for the final performance of a singer. Considering the importance of granulations, researchers proposed weighted generalized multi-granulation rough sets and some granulation weighted methods [29–31]. Nowadays, there are few studies on the combination of weighted generalized multi-granulation and DTRS. Ji et al. [30] investigated weighted multi-granulation intuitionistic fuzzy rough sets for decision-making. Guo et al. [29] proposed weighted generalized multi-granulation decision-theoretic rough sets based on the double weighted method for multi-source decision systems. The research of combining weighted generalized multi-granulation with interval-valued decision-theoretic rough sets for decision fusion is even less. It is meaningful to weight different granulations and then make decision fusion. Firstly, when the size of data sources is too large and some of them have to be deleted, weighted granulations can help to delete some data sources with smaller weights according to specific requirements. Secondly, when people do not want to lose any original information, weighted multi-granulation decision models by using valuable information from each source provide more reliable and comprehensive decision fusion methods. Therefore, the combination of weighted generalized multi-granulation and IVDTRS is a promising research direction. Such combined models are more flexible and adaptive to multi-source decision fusion with different needs.

It must be noted that the objects and attributes of data that need to be processed are increasing dramatically. In such data sets, the computations of approximations about the above rough set models are very time-consuming. Before making decision fusion, people can choose attribute reduction methods corresponding to actual data types to reduce the dimension of each granulation. Chen et al. [32] proposed efficient reduction methods based on sample pair selections for category data. Tsang et al. [33] and Yang et al. [34] proposed efficient incremental reduction algorithms based on fuzzy rough sets for real-valued data sets. Liu et al. [35] provided an efficient attribute reduction method of dynamic data from the incremental perspective. Yang et al. [36,37] proposed attribute reduction methods based on neighborhood rough sets and dominance-based rough sets for incomplete data. Wang et al. [38] and [39] studied attribute reductions of fuzzy rough sets characterized by distance measures and defined two novel measures used for attribute reductions of heterogeneous data, respectively. Considering that the diversity among different feature subsets helps to improve classification accuracy, Ślęzak et al. [40,41] studied feature subset ensembles based on attribute reductions of rough sets. The main research of this paper is to propose a weighted granulation method, and then make decision fusion of multi-granulation.

The existing weighted granulation methods [29,31] are mainly based on the upper and lower approximations of concepts. The time complexity of these weighted methods is positively related to the size of objects and attributes. In large data sets, these weighted methods are extremely time-consuming or even infeasible. Therefore, we propose a new granulation weighted method based on the machine learning method decision tree. Notice that the class imbalance of data will affect the performance of decision tree learning algorithm to some extent. For the processing of this type of data, please see literatures [42–45]. For multi-granulation fusion criteria, we choose the majority decision principle. Based on the proposed weighted method and the majority decision criteria, we propose weighted generalized multi-granulation interval-valued decision-theoretic rough sets (WGM-IVDTRS) for the decision fusion of multi-source decision systems with different attribute sets.

The main contributions of this paper are as follows: (1) From the perspective of machine learning, a granulation weighted method is proposed based on the classification accuracy of decision tree learning. (2) We propose three feasible and stable multi-granulation decision fusion models for multi-source decision systems concerned. (3) By controlling the information level parameter, we can solve the multi-source decision problems which require different classification precisions. (4) When decision makers can accept the tolerance ability of intervals, it is feasible to solve the minimum risk decision under multi-granulation through information optimization theory.

The rest of the paper is organized as follows. In Section 2, the preliminaries on interval-valued decision-theoretic rough sets, the geometry average interval sorting method and weighted generalized multi-granulation rough sets are briefly introduced. In Section 3, we first propose geometry average sorting interval-valued decision-theoretic rough sets (IVDTRSG) under single granulation. In Section 4, the basic form of the WGM-IVDTRS model, a new granulation weighted method based on the classification accuracy of decision tree learning and three types of the WGM-IVDTRS model are proposed. In Section 5, the importance, feasibility and effectiveness of the proposed weighted method are verified by comparing it with other models, and the applicability of three WGM-IVDTRS models is obtained by experimental results and analyses. Finally, Section 6 concludes the paper and elaborates on future studies.

2. Preliminaries

Basic concepts and notations of interval-valued decision-theoretic rough sets, the geometry average interval sorting method and weighted generalized multi-granulation rough sets are briefly reviewed in this section.

Table 1
An interval-valued loss function.

	X (P)	X^C (N)
a_P	$\widetilde{\lambda}_{PP}=[\lambda^-_{PP},\lambda^+_{PP}]$	$\widetilde{\lambda}_{PN}=[\lambda^-_{PN},\lambda^+_{PN}]$
a_B	$\widetilde{\lambda}_{BP}=[\lambda^-_{BP},\lambda^+_{BP}]$	$\widetilde{\lambda}_{BN}=[\lambda^-_{BN},\lambda^+_{BN}]$
a_N	$\widetilde{\lambda}_{NP}=[\lambda^-_{NP},\lambda^+_{NP}]$	$\widetilde{\lambda}_{NN}=[\lambda^-_{NN},\lambda^+_{NN}]$

2.1. Interval-valued decision-theoretic rough sets (IVDTRS) and two specific methods for determining parameters

In the interval-valued loss function, Liang et al. [16] proposed interval-valued decision-theoretic rough sets based on the Bayesian minimum risk decision. Firstly, we briefly introduce the basic content of the IVDTRS model.

Let $\Omega = \{X, X^C\}$ be a finite set of states indicating that a research object is in a concept X or not in X. $A = \{a_P, a_B, a_N\}$ be a set of three actions with respect to a state, where P, B and N represent the three actions in classifying objects, deciding positive region $pos(X)$, deciding boundary region $bnd(X)$ and deciding negative region $neg(X)$, respectively. When the values of a loss function are intervals, we call it an interval-valued loss function. Detailed descriptions are shown in Table 1.

In Table 1, $\widetilde{\lambda}_{PP}$, $\widetilde{\lambda}_{BP}$ and $\widetilde{\lambda}_{NP}$ denote the losses incurred for taking actions a_P, a_N and a_B, respectively, when an object belongs to X; and $\widetilde{\lambda}_{PN}$, $\widetilde{\lambda}_{BN}$ and $\widetilde{\lambda}_{NN}$ denote the losses incurred for taking the same actions when the object does not belong to X. It is true that the loss of $x \in pos(X)$ is smallest and the loss of $x \in pos(X)$ and $x \in bnd(X)$ are strictly smaller than the loss of $x \in neg(X)$ when $x \in X$, the reverse of the order of losses is used for $x \in \sim X$. Usually, there are $\lambda^-_{PP} \le \lambda^-_{BP} < \lambda^-_{NP}$, $\lambda^+_{PP} \le \lambda^+_{BP} < \lambda^+_{NP}$, $\lambda^-_{NN} \le \lambda^-_{BN} < \lambda^-_{PN}$, $\lambda^+_{NN} \le \lambda^+_{BN} < \lambda^+_{PN}$. Note that $x \in pos(X)$, $x \in bnd(X)$ and $x \in neg(X)$ mean that x is classified into positive region, boundary region and negative region, respectively.

For any object $x \in U$, the expected loss $R(\widetilde{a_i|[x]_R})$ $(i = P, B, N)$ associated with taking different actions can be expressed as

$$\begin{aligned}
R(\widetilde{a_P|[x]_R}) &= \widetilde{\lambda_{PP}}P(X|[x]_R) + \widetilde{\lambda_{PN}}P(X^C|[x]_R)\\
&= [\lambda^-_{PP}P(X|[x]_R)+\lambda^-_{PN}P(X^C|[x]_R),\ \lambda^+_{PP}P(X|[x]_R)+\lambda^+_{PN}P(X^C|[x]_R)],\\
R(\widetilde{a_B|[x]_R}) &= \widetilde{\lambda_{BP}}P(X|[x]_R) + \widetilde{\lambda_{BN}}P(X^C|[x]_R)\\
&= [\lambda^-_{BP}P(X|[x]_R)+\lambda^-_{BN}P(X^C|[x]_R),\ \lambda^+_{BP}P(X|[x]_R)+\lambda^+_{BN}P(X^C|[x]_R)],\\
R(\widetilde{a_N|[x]_R}) &= \widetilde{\lambda_{NP}}P(X|[x]_R) + \widetilde{\lambda_{NN}}P(X^C|[x]_R)\\
&= [\lambda^-_{NP}P(X|[x]_R)+\lambda^-_{NN}P(X^C|[x]_R),\ \lambda^+_{NP}P(X|[x]_R)+\lambda^+_{NN}P(X^C|[x]_R)].
\end{aligned} \tag{1}$$

where R is an equivalence relation generated by condition attributes, $[x]_R$ is the equivalence class containing x and $P(X|[x]_R) = |X \cap [x]_R|/|[x]_R|$. Based on the Bayesian decision procedure, the minimum-risk decision rules are

(P) If $R(\widetilde{a_P|[x]_R}) \le R(\widetilde{a_B|[x]_R})$ and $R(\widetilde{a_P|[x]_R}) \le R(\widetilde{a_N|[x]_R})$, decide $x \in pos(X)$;

(B) If $R(\widetilde{a_B|[x]_R}) \le R(\widetilde{a_P|[x]_R})$ and $R(\widetilde{a_B|[x]_R}) \le R(\widetilde{a_N|[x]_R})$, decide $x \in bnd(X)$;

(N) If $R(\widetilde{a_N|[x]_R}) \le R(\widetilde{a_B|[x]_R})$ and $R(\widetilde{a_N|[x]_R}) \le R(\widetilde{a_P|[x]_R})$, decide $x \in neg(X)$.

Obviously, the minimum Bayesian decision process of IVDTRS is essentially a time-consuming interval comparison. Next, we introduce two specific representative IVDTRS models, namely interval-valued decision-theoretic rough sets with a certain ranking method (IVDTRSC) and interval-valued decision-theoretic rough sets with an optimization method (IVDTRSO). They presented two methods for determining threshold parameters. More detailed instructions are found in literature [16].

In the IVDTRSC model, the minimum-risk decision rules by using the inequality simplification are

(P_C) If $P(X|[x]_R) \ge \alpha_C$ and $P(X|[x]_R) \ge \gamma_C$, decide $x \in pos(X)$;

(B_C) If $P(X|[x]_R) \le \alpha_C$ and $P(X|[x]_R) \ge \beta_C$, decide $x \in bnd(X)$;

(N_C) If $P(X|[x]_R) \le \beta_C$ and $P(X|[x]_R) \le \gamma_C$, decide $x \in neg(X)$;

where

$$\begin{aligned}
\alpha_C &= \frac{m_\theta(\widetilde{\lambda}_{PN}) - m_\theta(\widetilde{\lambda}_{BN})}{(m_\theta(\widetilde{\lambda}_{PN}) - m_\theta(\widetilde{\lambda}_{BN})) + (m_\theta(\widetilde{\lambda}_{BP}) - m_\theta(\widetilde{\lambda}_{PP}))},\\
\beta_C &= \frac{m_\theta(\widetilde{\lambda}_{BN}) - m_\theta(\widetilde{\lambda}_{NN})}{(m_\theta(\widetilde{\lambda}_{BN}) - m_\theta(\widetilde{\lambda}_{NN})) + (m_\theta(\widetilde{\lambda}_{NP}) - m_\theta(\widetilde{\lambda}_{BP}))},\\
\gamma_C &= \frac{m_\theta(\widetilde{\lambda}_{PN}) - m_\theta(\widetilde{\lambda}_{NN})}{(m_\theta(\widetilde{\lambda}_{PN}) - m_\theta(\widetilde{\lambda}_{NN})) + (m_\theta(\widetilde{\lambda}_{NP}) - m_\theta(\widetilde{\lambda}_{PP}))}.
\end{aligned} \tag{2}$$

Symbol $m_\theta(\widetilde{\lambda_\bullet})$ denotes the transformed formula of the interval $\widetilde{\lambda_\bullet} = [\lambda^-_\bullet, \lambda^+_\bullet]$, which can be calculated by $m_\theta(\widetilde{\lambda_\bullet}) = (1-\theta)\lambda^-_\bullet + \theta\lambda^+_\bullet$. Parameter θ is determined by the risk preferences of decision makers. The process of changing parameter θ from 0 to 1 reflects the risk preference attitude from optimism to pessimism.

When $\alpha_C > \beta_C$, then we have the following three-way decision rules:

(P_{C1}) If $P(X|[x]_R) \ge \alpha_C$, then decide $x \in pos(X)$;

(B_{C1}) If $\beta_C < P(X|[x]_R) < \alpha_C$, then decide $x \in bnd(X)$;

(N_{C1}) If $P(X|[x]_R) \le \beta_C$, then decide $x \in neg(X)$.

Two-way decision determined by γ_C in the condition $\alpha_C \le \beta_C$ is a special case of three-way decision. This paper mainly studies the interval-valued decision-theoretic rough set model under the general case, namely $\alpha_C > \beta_C$.

In the IVDTRSO model, the minimum-risk decision rules are

(P_O) If $P(X|[x]_R) \ge \alpha_O$ and $P(X|[x]_R) \ge \gamma_O$, decide $x \in pos(X)$;

(B_O) If $P(X|[x]_R) \le \alpha_O$ and $P(X|[x]_R) \ge \beta_O$, decide $x \in bnd(X)$;

(N_O) If $P(X|[x]_R) \le \beta_O$ and $P(X|[x]_R) \le \gamma_O$, decide $x \in neg(X)$;

where decision risk parameters $\alpha_O, \beta_O, \gamma_O$ are determined by means of the overall uncertainty minimization process. In case $\alpha_O > \beta_O$, the IVDTRSO model has the following decision rules:

(P_O) If $P(X|[x]_R) \ge \alpha_O$, then decide $x \in pos(X)$;

(B_O) If $\beta_O < P(X|[x]_R) < \alpha_O$, then decide $x \in bnd(X)$;

(N_O) If $P(X|[x]_R) \le \beta_O$, then decide $x \in neg(X)$.

The process of finding optimal parameters to minimize overall uncertainty is given as follows:

Let $\pi_{DT} = \{C, \sim C\}$ denote the partition of the universe U generated by the set DT of decision attributes, where DT divides U into two states. For the state C, a pair of thresholds (α_O, β_O) divides the universe into three regions corresponding to acceptance, delay, and rejection decisions, namely $pos_{(\alpha_O,\beta_O)}(C)$, $bnd_{(\alpha_O,\beta_O)}(C)$ and $neg_{(\alpha_O,\beta_O)}(C)$. Let $\pi_{(\alpha_O,\beta_O)}$ denote the partition of the universe generated by (α_O, β_O). The overall uncertainty of three regions is

$$\begin{aligned}
H(\pi_{DT}|\pi_{(\alpha_O,\beta_O)}) = {}& P(pos_{(\alpha_O,\beta_O)}(C))H(\pi_{DT}|pos_{(\alpha_O,\beta_O)}(C))\\
&+ P(bnd_{(\alpha_O,\beta_O)}(C))H(\pi_{DT}|bnd_{(\alpha_O,\beta_O)}(C))\\
&+ P(neg_{(\alpha_O,\beta_O)}(C))H(\pi_{DT}|neg_{(\alpha_O,\beta_O)}(C))
\end{aligned} \tag{3}$$

where $P(\triangle_{(\alpha_O,\beta_O)}(C)) = \frac{|\triangle_{(\alpha_O,\beta_O)}(C)|}{|U|}$ and $H(\pi_{DT}|\triangle_{(\alpha_O,\beta_O)}(C)) = -P(C|\triangle_{(\alpha_O,\beta_O)}(C))logP(C|\triangle_{(\alpha_O,\beta_O)}(C)) - P(\sim C|\triangle_{(\alpha_O,\beta_O)}(C))\ logP(\sim C|\triangle_{(\alpha_O,\beta_O)}(C))$, $(\triangle = pos, bnd, neg)$. Conditional probability $P(C|\triangle_{(\alpha_O,\beta_O)}(C))$ $(\triangle = pos, bnd, neg)$ can be calculated by $P(C|\triangle_{(\alpha_O,\beta_O)}(C)) = \frac{|\triangle_{(\alpha_O,\beta_O)}(C)\cap C|}{|\triangle_{(\alpha_O,\beta_O)}(C)|}$. The pair of optimal parameters is

$$\arg\min_{(\alpha_O,\beta_O)} \ H(\pi_{DT}|\pi_{(\alpha_O,\beta_O)})$$
$$s.t. = \begin{cases} \lambda_{PP}^- \le \lambda_{PP} \le \lambda_{PP}^+, \\ \lambda_{BP}^- \le \lambda_{BP} \le \lambda_{BP}^+, \\ \lambda_{NP}^- \le \lambda_{NP} \le \lambda_{NP}^+, \\ \lambda_{PN}^- \le \lambda_{PN} \le \lambda_{PN}^+, \\ \lambda_{BN}^- \le \lambda_{BN} \le \lambda_{BN}^+, \\ \lambda_{NN}^- \le \lambda_{NN} \le \lambda_{NN}^+, \\ \lambda_{PP} \le \lambda_{BP} \le \lambda_{NP}, \\ \lambda_{NN} \le \lambda_{BN} \le \lambda_{PN}, \end{cases} \tag{4}$$

where $\alpha_O = \frac{\lambda_{PN}-\lambda_{BN}}{(\lambda_{PN}-\lambda_{BN})+(\lambda_{BP}-\lambda_{PP})}$ and $\beta_O = \frac{\lambda_{BN}-\lambda_{NN}}{(\lambda_{BN}-\lambda_{NN})+(\lambda_{NP}-\lambda_{BP})}$.

2.2. The geometry average interval sorting method

Let $\widetilde{\lambda}_1, \widetilde{\lambda}_2, \ldots, \widetilde{\lambda}_n$ be n intervals, where $\widetilde{\lambda}_i = [\lambda_i^-, \lambda_i^+] = \{x \,|\lambda_i^- \le x \le \lambda_i^+\}$ and λ_i^-, λ_i^+ are real numbers. In the geometry average interval sorting method [17], the process of sorting n intervals is presented as follows:

Step 1 is to calculate the geometry average sorting function of interval $\widetilde{\lambda}_i$, namely $G_\eta(\widetilde{\lambda}_i) = (\lambda_i^-)^{1-\eta}(\lambda_i^+)^\eta$, $\eta \in [0,1]$ $(i = 1,2,\ldots,n)$.

Step 2 is to calculate the credibility of $\widetilde{\lambda}_i \ge \widetilde{\lambda}_j$ $(i,j = 1,2,\ldots,n)$ and establish the credibility matrix P, namely

$$P = \begin{pmatrix} 1 & \eta_{12} & \cdots & \eta_{1n} \\ \eta_{21} & 1 & \cdots & \eta_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \eta_{n1} & \eta_{n2} & \cdots & 1 \end{pmatrix}$$

where η_{ij} denotes the credibility of $\widetilde{\lambda}_i \ge \widetilde{\lambda}_j$ $(i \ne j)$ and its specific expression is

$$\eta_{ij} = \begin{cases} 1 - inf\{\eta|G_\eta(\widetilde{\lambda}_i) \ge G_\eta(\widetilde{\lambda}_j)\}, & \text{if } inf\{\eta|G_\eta(\widetilde{\lambda}_i) \ge G_\eta(\widetilde{\lambda}_j)\} > 0; \\ sup\{\eta|G_\eta(\widetilde{\lambda}_i) \ge G_\eta(\widetilde{\lambda}_j)\}, & \text{if } inf\{\eta|G_\eta(\widetilde{\lambda}_i) \ge G_\eta(\widetilde{\lambda}_j)\} = 0. \end{cases} \tag{5}$$

Step 3 is to sort intervals according to the credibility matrix. If there are $n+1-q$ $(q = 1,2,\ldots,n)$ elements greater than 0.5 in the ith row, then the ith interval was ranked qth.

For example, given two intervals $\widetilde{\lambda}_1 = [0.1, 0.5]$, $\widetilde{\lambda}_2 = [0.2, 0.4]$, when $0.756 \le \eta \le 1$, there is $G_\eta(\widetilde{\lambda}_1) \ge G_\eta(\widetilde{\lambda}_2)$. Therefore, $inf\{\eta \mid G_\eta(\widetilde{\lambda}_1) \ge G_\eta(\widetilde{\lambda}_2)\} = 0.756 > 0$. So $\eta_{12} = 1 - inf\{\eta \mid G_\eta(\widetilde{\lambda}_1) \ge G_\eta(\widetilde{\lambda}_2)\} = 0.244 < 0.5$. Therefore, according to geometry average sorting method, we have $\widetilde{\lambda}_2 \ge \widetilde{\lambda}_1$.

2.3. Weighted generalized multi-granulation rough sets (WGMRS)

First, we briefly introduce the concept of information systems. Let $I = (U, AT \cup DT, V, F)$ be an information system, where universe U is a nonempty finite set of objects; AT is a set of condition attributes and DT is a set of decision attributes; $V = \cup_{a\in AT\cup DT} V_a$ is the domain of attribute values; $F : U \times \{AT \cup DT\} \to V$ is an information function, i.e., $\forall a \in AT \cup DT$, $x \in U$, that $F(x,a) \in V_a$ is the value of object x about attribute a. Usually, we use $U/DT = \{D_1, D_2, \ldots, D_t\}$ to denote the partition of U under DT, where D_i consists of objects with the same description of decision attribute.

In information system $I = (U, AT \cup DT, V, F)$, $\forall A_i \subseteq AT$, $i = 1,2,\ldots,s$, $s \le 2^{AT}$, for any subset X of U, the weighted generalized multi-granulation upper and lower approximations of X with respect to $\sum_{i=1}^s A_i$ in literature [31] are defined as

$$\begin{aligned} &\overline{GM}^w_{\sum_{i=1}^s A_i}(X) = \{x \in U|\, (\sum_{i=1}^s \varpi_i(1 - S^{A_i}_{\sim X}(x))) > 1 - \varphi\}, \\ &\underline{GM}^w_{\sum_{i=1}^s A_i}(X) = \{x \in U|\, (\sum_{i=1}^s \varpi_i S^{A_i}_X(x)) \ge \varphi\}, \end{aligned} \tag{6}$$

where the superscript w is usually used to distinguish the approximations of the weighted generalized multi-granulation rough set model and the generalized multi-granulation rough set model. Parameter ϖ_i $(i = 1,2,\ldots,s)$ denotes the weight of granulation A_i, and $\sum_{i=1}^s \varpi_i = 1$. $S^{A_i}_X(x)$ denotes the support characteristic function of x with respect to concept X under A_i. Let $[x]_{A_i}$ denote the equivalence class of x under granulation A_i, which consists of objects with the same description as x under A_i. If $[x]_{A_i} \subseteq X$, then $S^{A_i}_X(x) = 1$; otherwise $S^{A_i}_X(x) = 0$. Parameter $\varphi \in (0.5, 1]$ denotes the information level with respect to $\sum_{i=1}^s A_i$. In the WGMRS model, positive, boundary and negative regions of X are defined as

$$\begin{aligned} &pos(X) = \underline{GM}^w_{\sum_{i=1}^s A_i}(X); \\ &bnd(X) = \overline{GM}^w_{\sum_{i=1}^s A_i}(X) - \underline{GM}^w_{\sum_{i=1}^s A_i}(X); \\ &neg(X) = \sim \overline{GM}^w_{\sum_{i=1}^s A_i}(X). \end{aligned} \tag{7}$$

In addition, we briefly introduce three representative weighted methods, namely approximation accuracy, approximation quality and granulation entropy weighted methods. For more weighted granulation methods, please refer to literature [31].

Let $I = (U, AT \cup DT, V, F)$ be an information system, $A_i \subseteq AT$ $(i = 1,2,\ldots,s)$ and $U/DT = \{D_1, D_2, \ldots, D_t\}$. The approximation accuracy weight of granulation A_i is defined as

$$w_i = \frac{\sum_{j=1}^t \alpha_i(D_j)}{\sum_{i=1}^s \sum_{j=1}^t \alpha_i(D_j)}, \tag{8}$$

where $\alpha_i(D_j) = \frac{|\underline{A_i}(D_j)|}{|\overline{A_i}(D_j)|}$ is the approximation accuracy of A_i about DT. $\overline{A_i}(D_j)$ and $\underline{A_i}(D_j)$ represent the Pawlak upper and lower approximations of D_j $(j = 1,2,\ldots,t)$ under A_i, respectively. The approximation quality weight of granulation A_i is defined as

$$w_i = \frac{a(A_i)}{\sum_{i=1}^s a(A_i)}, \tag{9}$$

where $a(A_i) = \frac{\sum_{j=1}^t |\underline{A_i}(D_j)|}{\sum_{j=1}^t |\overline{A_i}(D_j)|}$ is the approximation quality of A_i about DT. The granulation entropy weight of granulation A_i is defined as

$$w_i = \frac{\sum_{j=1}^t E_{A_i}(D_j)}{\sum_{i=1}^s \sum_{j=1}^t E_{A_i}(D_j)}, \tag{10}$$

where $E_{A_i}(D_j)$ is the rough entropy of D_j about A_i, which can be calculated by $E_{A_i}(D_j) = \rho_i(D_j)E(A_i) = (1 - \alpha_i(D_j))(-\sum_{k=1}^{|U|} \frac{|[x_k]_{A_i}|}{|U|} * log_2 \frac{|[x_k]_{A_i}|}{|U|})$, where $|\cdot|$ denotes the cardinality of a set.

3. Geometry average sorting interval-valued decision-theoretic rough sets (IVDTRSG)

Decision-theoretic rough set (DTRS) theory based on Bayesian minimum decision risk provides a reasonable semantic interpretation for decision-making process. The determination of threshold parameters mainly relies on the loss function. Considering intervals can reflect a certain tolerant ability during the evaluation and the Bayesian minimum decision is actually the process of sorting intervals in the interval-valued decision-theoretic rough set (IVDRTS) method, we propose a new IVDTRS model based on geometry average interval sorting method in this section, namely geometry average sorting interval-valued decision-theoretic rough sets (IVDTRSG).

Firstly, the basic framework of the IVDTRSG model is described.

Let $\Omega = \{C, \sim C\}$ be a state set and $A = \{a_P, a_B, a_N\}$ be an action set, where two state sets are complementary decision classes and a_P, a_B, a_N denote accepting, delaying and rejecting decisions, respectively. The loss function usually satisfies conditions $\lambda^-_{PP} \le \lambda^-_{BP} < \lambda^-_{NP}$, $\lambda^+_{PP} \le \lambda^+_{BP} < \lambda^+_{NP}$, $\lambda^-_{NN} \le \lambda^-_{BN} < \lambda^-_{PN}$, $\lambda^+_{NN} \le \lambda^+_{BN} < \lambda^+_{PN}$. The expected loss of any object x ($x \in U$) associated with three actions in IVDTRSG model are shown in formula (1). For each object, by sorting intervals, we find the action causing the smallest expected loss. Similar to the certain ranking method and the optimization method, we use geometry average method to sort the intervals.

According to conditions $\lambda^-_{PP} \le \lambda^-_{BP} < \lambda^-_{NP}, \lambda^+_{PP} \le \lambda^+_{BP} < \lambda^+_{NP}$, $\lambda^-_{NN} \le \lambda^-_{BN} < \lambda^-_{PN}$, $\lambda^+_{NN} \le \lambda^+_{BN} < \lambda^+_{PN}$ and $G_\eta(\widetilde{\lambda_i}) = (\lambda^-_i)^{1-\eta}(\lambda^+_i)^\eta$, there are $G_\eta(\widetilde{\lambda}_{PP}) \le G_\eta(\widetilde{\lambda}_{BP}) < G_\eta(\widetilde{\lambda}_{NP})$ and $G_\eta(\widetilde{\lambda}_{NN}) \le G_\eta(\widetilde{\lambda}_{BN}) < G_\eta(\widetilde{\lambda}_{PN})$. Then the minimum-risk decision rules are expressed equivalently as follows:

(P_G) If $G_\eta(\widetilde{R(a_P|[x]_R)}) \le G_\eta(\widetilde{R(a_B|[x]_R)})$ and $G_\eta(\widetilde{R(a_P|[x]_R)}) \le G_\eta(\widetilde{R(a_N|[x]_R)})$, then decide $x \in pos(X)$;

(B_G) If $G_\eta(\widetilde{R(a_B|[x]_R)}) \le G_\eta(\widetilde{R(a_P|[x]_R)})$ and $G_\eta(\widetilde{R(a_B|[x]_R)}) \le G_\eta(\widetilde{R(a_N|[x]_R)})$, then decide $x \in bnd(X)$;

(N_G) If $G_\eta(\widetilde{R(a_N|[x]_R)}) \le G_\eta(\widetilde{R(a_P|[x]_R)})$ and $G_\eta(\widetilde{R(a_N|[x]_R)}) \le G_\eta(\widetilde{R(a_B|[x]_R)})$, then decide $x \in neg(X)$; where $G_\eta(\widetilde{R(a_\bullet|[x]_R)}) = [\lambda^-_{\bullet P}P(X|[x]_R) + \lambda^-_{\bullet N}(1 - P(X|[x]_R))]^{1-\eta}[\lambda^+_{\bullet P}P(X|[x]_R) + \lambda^+_{\bullet N}(1 - P(X|[x]_R))]^\eta$ ($\bullet = P, B, N$).

By simplification, we have the following decision rules:

(P_G) If $P(X|[x]_R) \ge \alpha_G$ and $P(X|[x]_R) \ge \gamma_G$, then decide $x \in pos(X)$;

(B_G) If $P(X|[x]_R) \le \alpha_G$ and $P(X|[x]_R) \ge \beta_G$, then decide $x \in bnd(X)$;

(N_G) If $P(X|[x]_R) \le \beta_G$ and $P(X|[x]_R) \le \gamma_G$, then decide $x \in neg(X)$; where

$$\begin{aligned}
\alpha_G &= \frac{G_\eta(\widetilde{\lambda}_{PN}) - G_\eta(\widetilde{\lambda}_{BN})}{G_\eta(\widetilde{\lambda}_{PN}) - G_\eta(\widetilde{\lambda}_{BN}) + G_\eta(\widetilde{\lambda}_{BP}) - G_\eta(\widetilde{\lambda}_{PP})},\\
\beta_G &= \frac{G_\eta(\widetilde{\lambda}_{BN}) - G_\eta(\widetilde{\lambda}_{NN})}{G_\eta(\widetilde{\lambda}_{BN}) - G_\eta(\widetilde{\lambda}_{NN}) + G_\eta(\widetilde{\lambda}_{NP}) - G_\eta(\widetilde{\lambda}_{BP})},\\
\gamma_G &= \frac{G_\eta(\widetilde{\lambda}_{PN}) - G_\eta(\widetilde{\lambda}_{NN})}{G_\eta(\widetilde{\lambda}_{PN}) - G_\eta(\widetilde{\lambda}_{NN}) + G_\eta(\widetilde{\lambda}_{NP}) - G_\eta(\widetilde{\lambda}_{PP})}.
\end{aligned} \tag{11}$$

If a loss function further satisfies the condition: $\frac{G_\eta(\widetilde{\lambda}_{BP})-G_\eta(\widetilde{\lambda}_{PP})}{G_\eta(\widetilde{\lambda}_{PN})-G_\eta(\widetilde{\lambda}_{BN})} \le \frac{G_\eta(\widetilde{\lambda}_{NP})-G_\eta(\widetilde{\lambda}_{BP})}{G_\eta(\widetilde{\lambda}_{BN})-G_\eta(\widetilde{\lambda}_{NN})}$ then we can get $\alpha_G \ge \gamma_G \ge \beta_G$. In this paper, we mainly study the decision loss function satisfying this condition. In this condition, the above decision rules can be further simplified as

(P_G) If $P(X|[x]_R) \ge \alpha_G$, then decide $x \in pos(X)$;

(B_G) If $\beta_G < P(X|[x]_R) < \alpha_G$, then decide $x \in bnd(X)$;

(N_G) If $P(X|[x]_R) \le \beta_G$, then decide $x \in neg(X)$.

Then the upper and lower approximations of the IVDTRSG model are expressed as

$$\begin{aligned}
&\overline{R}_{(\alpha_G,\beta_G)}(X) = \{x \in U | P(X|[x]_R) > \beta_G\};\\
&\underline{R}_{(\alpha_G,\beta_G)}(X) = \{x \in U | P(X|[x]_R) \ge \alpha_G\}.
\end{aligned} \tag{12}$$

In the following, we illustrate the similarities and differences between IVDTRSG and IVDTRS by an example. For the convenience of comparison, data information and the interval-valued loss function are derived from literature [46] and [16], respectively.

Example 1. Let C be a concept and E_i ($i \in \{1, 2, \ldots, 15\}$) denote equivalence class generated by conditional attributes in universe. Detailed statistics of probability information [46] are shown in Table 2. The values of loss function [16] are $\widetilde{\lambda}_{PP} = [1.5334, 3.1545]$, $\widetilde{\lambda}_{PN} = [3.3798, 5.7363]$, $\widetilde{\lambda}_{BP} = [2.4087, 3.3474]$, $\widetilde{\lambda}_{BN} = [2.4496, 3.3673]$, $\widetilde{\lambda}_{NP} = [3.2707, 4.5747]$, $\widetilde{\lambda}_{NN} = [0.7668, 3.1200]$. Assuming that parameter η determined by decision makers is 0.6.

In the IVDTRS model, for each equivalence class, we first calculate the expected losses about three actions. Detailed results are shown in Table 3.

Then for each equivalence class, the action corresponding to the minimum expected loss is found based on different interval ranking methods such as θ ranking method and the optimization method [16]. In these two methods, the threshold parameters (α, β) are obtained as (0.7554, 0.4802) and (0.8211, 0.3833), respectively. Furthermore, three rough regions and two approximations of concept C can be obtained.

In the IVDTRSG model, we first calculate the values of geometry average sorting function, namely $G_\eta(\widetilde{\lambda}_{PP}) = 2.3639$, $G_\eta(\widetilde{\lambda}_{PN}) = 4.6423$, $G_\eta(\widetilde{\lambda}_{BP}) = 2.9345$, $G_\eta(\widetilde{\lambda}_{BN}) = 2.9649$, $G_\eta(\widetilde{\lambda}_{NP}) = 4.0001$ and $G_\eta(\widetilde{\lambda}_{NN}) = 1.7798$. Then threshold parameters α_G and β_G can be known as $\alpha_G = 0.7462$ and $\beta_G = 0.5265$. According to the probability information, we can directly get three regions as follows: $pos(X) = E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5 \cup E_6 \cup E_7$, $bnd(X) = E_8$ and $neg(X) = E_9 \cup E_{10} \cup E_{11} \cup E_{12} \cup E_{13} \cup E_{14} \cup E_{15}$. The upper and lower approximations of concept C in the IVDTRSG model can also be obtained, namely $\overline{R}_{(\alpha_G,\beta_G)}(X) = E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5 \cup E_6 \cup E_7$ and $\underline{R}_{(\alpha_G,\beta_G)}(X) = E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5 \cup E_6 \cup E_7 \cup E_8$.

Through comparative analysis, we get the similarities and differences between IVDTRSG and IVDTRS. The similarities of IVDTRSG and IVDTRS are as follows:

• The loss function and the probability information between equivalence classes and concepts are two key factors of IVDTRSG and IVDTRS in decision-making. Both IVDTRSG and IVDTRS consider the interval-valued loss function.

• The decision criteria of IVDTRSG and IVDTRS are all Bayesian minimum decision theory. These two models provide methods for the three-way decisions under interval-valued loss environment.

• IVDTRSG and IVDTRS are generalizations of decision-theoretic rough sets (DTRS). When the values of the loss function become real numbers, both IVDTRSG and IVDTRS degenerate into DTRS.

The differences of IVDTRSG and IVDTRS are as follows:

• The IVDTRS model needs to compare intervals while the IVDTRSG model compares real numbers directly when searching the action corresponding to the minimum expected loss for each equivalence classes.

• The IVDTRS model provides a basic framework for decision-making in interval-valued loss function. The IVDTRSG model is a specific decision method by transforming intervals to precise values in the case of acceptable risk. According to different requirements, researchers introduce different interval ranking methods into IVDTRS, such as representative certain methods [47] estimating a precise value for the interval, degree of possibility ranking methods [48] for comparing intervals and optimization methods [49] based on different objective functions. Then many specific IVDTRS models are obtained such as IVDTRSC and IVDTRSO.

4. Weighted generalized multi-granulation interval-valued decision-theoretic rough sets (WGM-IVDTRS)

Considering intervals can reflect a certain tolerant ability during the evaluation and different granulations may not be equally important in practical problems, we propose the general form of

Table 2
Probabilistic information.

	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}	E_{11}	E_{12}	E_{13}	E_{14}	E_{15}
$P(E_i)$	0.0177	0.1285	0.0137	0.1352	0.0580	0.0069	0.0498	0.1070	0.1155	0.0792	0.0998	0.1299	0.0080	0.0441	0.0067
$P(C\|E_i)$	1.0	1.0	1.0	1.0	0.9	0.8	0.8	0.6	0.5	0.4	0.4	0.2	0.1	0.0	0.0

Table 3
The expected losses.

	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8
R_P	[1.5334, 3.1545]	[1.5334, 3.1545]	[1.5334, 3.1545]	[1.5334, 3.1545]	[1.7180, 3.4127]	[1.9027, 3.6709]	[1.9027, 3.6709]	[2.2720, 4.1872]
R_B	[2.4087, 3.3474]	[2.4087, 3.3474]	[2.4087, 3.3474]	[2.4087, 3.3474]	[2.4128, 3.3494]	[2.4169, 3.3514]	[2.4169, 3.3514]	[2.4251, 3.3554]
R_N	[3.2707, 4.5747]	[3.2707, 4.5747]	[3.2707, 4.5747]	[3.2707, 4.5747]	[3.0203, 4.4292]	[2.7699, 4.2838]	[2.7699, 4.2838]	[2.2691, 3.9928]
	E_9	E_{10}	E_{11}	E_{12}	E_{13}	E_{14}	E_{15}	
R_P	[2.4566, 4.4454]	[2.6412, 4.7036]	[2.6412, 4.7036]	[3.0105, 5.2199]	[3.1952, 5.4781]	[3.3798, 5.7363]	[3.3798, 5.7363]	
R_B	[2.4291, 3.3574]	[2.4332, 3.3593]	[2.4332, 3.3593]	[2.4414, 3.3633]	[2.4455, 3.3653]	[2.4496, 3.3673]	[2.4496, 3.3673]	
R_N	[2.0188, 3.8474]	[1.7684, 3.7019]	[1.7684, 3.7019]	[1.2676, 3.4109]	[1.0172, 3.2655]	[0.7668, 3.1200]	[0.7668, 3.1200]	

weighted generalized multi-granulation interval-valued decision-theoretic rough sets (WGM-IVDTRS) based on the classification accuracy of decision tree learning in this section. Based on different methods of determining threshold parameters under an interval-valued loss function, three types of the WGM-IVDTRS model are constructed. The first WGM-IVDTRS model is weighted generalized multi-granulation certain ranking interval-valued decision-theoretic rough sets (WGM-IVDTRSC), in which parameter θ in the transformation formula can reflect the risk preferences of different decision makers. The second model is weighted generalized multi-granulation geometry average sorting interval-valued decision-theoretic rough sets (WGM-IVDTRSG), which in the process of converting intervals into real numbers, the geometry average sorting function has good performance. The third model is weighted generalized multi-granulation optimization interval-valued decision-theoretic rough sets (WGM-IVDTRSO), which can objectively reflect the uncertainty of information from the perspective of machine learning.

4.1. The general form of WGM-IVDTRS

Definition 4.1. Let $I = (U, AT \cup DT, V, F)$ be an information system, $A_i \subseteq AT$, $i = 1, 2, \ldots, s$ $(s \leq 2^{AT})$, $\varphi \in (0.5, 1]$. For any $X \subseteq U$, the weighted generalized multi-granulation upper and lower approximations of X with respect to $\sum_{i=1}^{s} A_i$ in the WGM-IVDTRS model are defined as

$$\begin{aligned} &\overline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | \sum_{i=1}^{s} WUS_X^{A_i}(x) > 1 - \varphi\}, \\ &\underline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | \sum_{i=1}^{s} WLS_X^{A_i}(x) \geq \varphi\}, \end{aligned} \tag{13}$$

where ϖ_i is the weight of granulation A_i, and $\sum_{i=1}^{s} \varpi_i = 1$. Parameter φ is the information level with respect to $\sum_{i=1}^{s} A_i$. Symbols $WUS_X^{A_i}(x)$ and $WLS_X^{A_i}(x)$ denote the weighted upper and lower support characteristic functions of x with respect to concept X under granulation A_i, respectively. For any x $(x \in U)$, if and only if $P(X|[x]_{A_i}) > \beta_i$, we have $WUS_X^{A_i}(x) = \varpi_i$; otherwise $WUS_X^{A_i}(x) = 0$. Similarly, if and only if $P(X|[x]_{A_i}) \geq \alpha_i$, we have $WLS_X^{A_i}(x) = \varpi_i$; otherwise $WLS_X^{A_i}(x) = 0$. Information level φ reflects people's risk preferences under multiple granulations. With the increase of φ, the degree of optimism reflected decreases.

According to Definition 4.1, if the sum of weighted lower support characteristic functions of object x with respect to concept X under all granulations is not smaller than φ, then object x definitely belong to X, namely $x \in \underline{WGM}_{\sum_{i=1}^{s} A_i}(X)$. If the sum of weighted upper support characteristic functions of object x with respect to X under all granulations is greater than $1 - \varphi$, then x probably belong to X, namely $x \in \overline{WGM}_{\sum_{i=1}^{s} A_i}(X)$.

In the WGM-IVDTRS model, positive, negative and boundary regions are defined as follows:

$$\begin{aligned} &pos(X) = \underline{WGM}_{\sum_{i=1}^{s} A_i}(X), \\ &neg(X) = \sim \overline{WGM}_{\sum_{i=1}^{s} A_i}(X), \\ &bnd(X) = \overline{WGM}_{\sum_{i=1}^{s} A_i}(X) - \underline{WGM}_{\sum_{i=1}^{s} A_i}(X). \end{aligned} \tag{14}$$

Considering two extreme cases of granulation requirements namely optimism and pessimism, we study weighted optimistic and pessimistic multi-granulation interval-valued decision-theoretic rough sets (WOM-IVDTRS and WPM-IVDTRS). Details are defined as follows.

Definition 4.2. Let $I = (U, AT \cup DT, V, F)$ be an information system, $A_i \subseteq AT$, $i = 1, 2, \ldots, s$ $(s \leq 2^{AT})$, $\varphi \in (0.5, 1]$. For any $X \subseteq U$, the weighted optimistic and pessimistic multi-granulation upper and lower approximations of X with respect to $\sum_{i=1}^{s} A_i$ in the WOM-IVDTRS and WPM-IVDTRS models are defined as

$$\begin{aligned} \overline{WOM}_{\sum_{i=1}^{s} A_i}(X) &= \{x \in U | \sum_{i=1}^{s} WUS_X^{A_i}(x) \geq 1\} \\ &= \{x \in U | \wedge_{i=1}^{s} \{P(X|[x]_{A_i}) > \beta_i\}\}, \\ \underline{WOM}_{\sum_{i=1}^{s} A_i}(X) &= \{x \in U | \sum_{i=1}^{s} WLS_X^{A_i}(x) > 0\} \\ &= \{x \in U | \vee_{i=1}^{s} \{P(X|[x]_{A_i}) \geq \alpha_i\}\}. \end{aligned} \tag{15}$$

$$\begin{aligned} \overline{WPM}_{\sum_{i=1}^{s} A_i}(X) &= \{x \in U | \sum_{i=1}^{s} WUS_X^{A_i}(x) > 0\} \\ &= \{x \in U | \vee_{i=1}^{s} \{P(X|[x]_{A_i}) > \beta_i\}\}, \\ \underline{WPM}_{\sum_{i=1}^{s} A_i}(X) &= \{x \in U | \sum_{i=1}^{s} WLS_X^{A_i}(x) \geq 1\} \\ &= \{x \in U | \wedge_{i=1}^{s} \{P(X|[x]_{A_i}) \geq \alpha_i\}\}. \end{aligned} \tag{16}$$

According to Definition 4.2, we give semantic explanations from the perspectives of support characteristic functions. From formula (15), if there is one granulation that makes the value of the weighted lower support characteristic function of x with respect to X greater than 0, then x definitely belong to X under optimistic risk preference, namely $x \in \underline{WOM}_{\sum_{i=1}^{s} A_i}(X)$; if the values of weighted upper support characteristic functions of x with respect to X under all granulations are all greater than 0, then x probably belong to X under optimistic risk preference, namely $x \in \overline{WOM}_{\sum_{i=1}^{s} A_i}(X)$. From formula (16), if the values of weighted lower support characteristic functions of object x with respect to X under all granulations are all greater than 0, then x definitely belong to X under pessimistic risk preference, namely $x \in \underline{WPM}_{\sum_{i=1}^{s} A_i}(X)$; if there is one granulation that makes the value of the weighted upper support characteristic function of x with respect to X greater than 0, then x probably belong to X under pessimistic risk preference, namely $x \in \underline{WPM}_{\sum_{i=1}^{s} A_i}(X)$.

Definition 4.3. Let $I = (U, AT \cup DT, V, F)$ be an information system, and $U/DT = \{D_1, D_2, \ldots, D_t\}$, where D_j $(j = 1, 2, \ldots, t)$ denotes a decision class. Then the approximation accuracy $acc(D_j)$ and classification error rate e of D_j in the WGM-IVDTRS model are defined as

$$acc(D_j) = \frac{|\underline{WGM}_{\sum_{i=1}^{s} A_i}(D_j)|}{|\overline{WGM}_{\sum_{i=1}^{s} A_i}(D_j)|}, \tag{17}$$

$$e = \frac{|D_j \cap neg(D_j)| + |\sim D_j \cap pos(D_j)|}{|U|}. \tag{18}$$

From Definition 4.3, the approximation accuracy of a decision class D_j is equal to the ratio of the cardinality of the lower approximation of this decision class to that of the upper approximation. The classification error rate of D_j is the percentage of objects in the universe, in which these objects mainly include objects belonging to D_j that are classified into negative region $neg(D_j)$ and objects that do not belong to D_j are classified into positive region $pos(D_j)$.

Considering the relationships among weighted generalized, optimistic and pessimistic multi-granulation and the relationship between multiple granular structures and single granular structure, we obtained the following conclusions.

Proposition 4.1. *Let $I = (U, AT \cup DT, V, F)$ be an information system, $A_i \subseteq AT$, $i = 1, 2, \ldots, s$ $(s \le 2^{AT})$, $\varphi \in (0.5, 1]$. For any $X \subseteq U$, the following conclusions hold:*

(1) $\overline{WPM}_{\sum_{i=1}^{s} A_i}(X) \subseteq \overline{WGM}_{\sum_{i=1}^{s} A_i}(X) \subseteq \overline{WOM}_{\sum_{i=1}^{s} A_i}(X)$;
$\underline{WOM}_{\sum_{i=1}^{s} A_i}(X) \subseteq \underline{WGM}_{\sum_{i=1}^{s} A_i}(X) \subseteq \underline{WPM}_{\sum_{i=1}^{s} A_i}(X)$.

(2) $\overline{WOM}_{\sum_{i=1}^{s} A_i}(X) \subseteq \overline{A_i}_{(\alpha_i,\beta_i)}(X)$, $\underline{WOM}_{\sum_{i=1}^{s} A_i}(X) \supseteq \underline{A_i}_{(\alpha_i,\beta_i)}(X)$;
$\overline{WPM}_{\sum_{i=1}^{s} A_i}(X) \supseteq \overline{A_i}_{(\alpha_i,\beta_i)}(X)$, $\underline{WPM}_{\sum_{i=1}^{s} A_i}(X) \subseteq \underline{A_i}_{(\alpha_i,\beta_i)}(X)$.

(3) $\overline{WOM}_{\sum_{i=1}^{s} A_i}(X) = \cap_{i=1}^{s} \overline{A_i}_{(\alpha_i,\beta_i)}(X)$, $\underline{WOM}_{\sum_{i=1}^{s} A_i}(X) = \cup_{i=1}^{s} \underline{A_i}_{(\alpha_i,\beta_i)}(X)$;
$\overline{WPM}_{\sum_{i=1}^{s} A_i}(X) = \cup_{i=1}^{s} \overline{A_i}_{(\alpha_i,\beta_i)}(X)$, $\underline{WPM}_{\sum_{i=1}^{s} A_i}(X) = \cap_{i=1}^{s} \underline{A_i}_{(\alpha_i,\beta_i)}(X)$.

In Proposition 4.1, from conclusion (1), we find that both the weighted pessimistic upper (optimistic lower) approximation of a concept and its weighted generalized upper (generalized lower) approximation are contained in its weighted optimistic upper (pessimistic lower) approximation. Moveover, the weighted pessimistic upper (optimistic lower) approximation of this concept is contained in its weighted generalized upper (generalized lower) approximation. By conclusion (2), the weighted optimistic upper (pessimistic lower) approximation of a concept is contained in the upper (lower) approximation of this concept under arbitrary granulation, and the weighted optimistic lower (pessimistic upper) approximation of a concept contains the lower (upper) approximation of this concept under arbitrary granulation. By conclusion (3), the weighted optimistic upper (pessimistic lower) approximation of a concept is the intersection of upper (lower) approximations of this concept under all granulations. The weighted optimistic lower (pessimistic upper) approximation of a concept is the union of lower (upper) approximations of this concept under all granulations.

The relationship between the WGM-IVDTRS model and the weighted generalized multi-granulation rough set (WGMRS) model are explored in the following.

Proposition 4.2. *Let $I = (U, AT \cup DT, V, F)$ be an information system, $A_i \subseteq AT$, $i = 1, 2, \ldots, s$ $(s \le 2^{AT})$, $\varphi \in (0.5, 1]$. For any $X \subseteq U$, based on certain constraints, the following conclusions hold:*

(1) *When $\varpi_1 = \varpi_2 = \cdots = \varpi_s$, there are*
$\overline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | \sum_{i=1}^{s} US_X^{A_i}(x)/s > 1 - \varphi\} \subseteq \overline{GM}^w_{\sum_{i=1}^{s} A_i}(X)$;
$\underline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | \sum_{i=1}^{s} LS_X^{A_i}(x)/s \ge \varphi\} \supseteq \underline{GM}^w_{\sum_{i=1}^{s} A_i}(X)$.

(2) *When $\varpi_i = 1 (i \in \{1, 2, \ldots, s\})$, there are*
$\overline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | WUS_X^{A_i}(x) = 1\} = \overline{A}_{i(\alpha_i,\beta_i)}(X) \subseteq \overline{GM}^w_{\sum_{i=1}^{s} A_i}(X)$;
$\underline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | WLS_X^{A_i}(x) = 1\} = \underline{A_i}_{(\alpha_i,\beta_i)}(X) \supseteq \underline{GM}^w_{\sum_{i=1}^{s} A_i}(X)$.

(3) *When $\alpha_1 = \alpha_2 = \cdots = \alpha_s = 1, \beta_1 = \beta_2 = \cdots = \beta_s = 0$, there are*
$\overline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | \sum_{i=1}^{s} \varpi_i S_X^{A_i}(x) > 1 - \varphi\} = \overline{GM}^w_{\sum_{i=1}^{s} A_i}(X)$;
$\underline{WGM}_{\sum_{i=1}^{s} A_i} = \{x \in U | \sum_{i=1}^{s} \varpi_i S_X^{A_i}(x) \ge \varphi\} = \underline{GM}^w_{\sum_{i=1}^{s} A_i}(X)$.

(4) *For $0 \le \beta_i < \alpha_i \le 1 (i = 1, 2, \ldots, s)$, there are*
$\overline{WGM}_{\sum_{i=1}^{s} A_i}(X) \subseteq \overline{GM}^w_{\sum_{i=1}^{s} A_i}(X)$; $\underline{WGM}_{\sum_{i=1}^{s} A_i} \supseteq \underline{GM}^w_{\sum_{i=1}^{s} A_i}(X)$.

Functions $US_X^{A_i}(x)$ and $LS_X^{A_i}(x)$ are the upper and lower support characteristic functions of $x \in U$ with respect to concept X under A_i. If $P(X|[x]_{A_i}) > \beta_i$, then $US_X^{A_i}(x) = 1$; otherwise, $US_X^{A_i}(x) = 0$. Similarly, if $P(X|[x]_{A_i}) \ge \alpha_i$, then $LS_X^{A_i}(x) = 1$; otherwise, $LS_X^{A_i}(x) = 0$.

According to the conclusions (1–4) of Proposition 4.2, it does not matter whether the importance of granulations is equal or not, the WGM-IVDTRS model is a generalized model of the WGMRS model. In particular, when $\alpha_1 = \alpha_2 = \cdots = \alpha_s = 1, \beta_1 = \beta_2 = \cdots = \beta_s = 0$, the WGM-IVDTRS model and the WGMRS model are equivalent. Moreover, it is obvious that the approximation accuracy of decision class X in the WGM-IVDTRS model is greater than or equal to the approximation accuracy of X in the WGMRS model. When $\varpi_i = 1$ $(i \in \{1, 2, \ldots, s\})$, the WGM-IVDTRS model will degenerate into the classical IVDTRS model under the single granular structure A_i.

Proposition 4.3. *Let $I = (U, AT \cup DT, V, F)$ be an information system, $U/DT = \{D_1, D_2, \ldots, D_t\}$, $A_i \subseteq AT$, $i = 1, 2, \ldots, s$ $(s \le 2^{AT})$ and $\varphi_1 \le \varphi_2 \in (0.5, 1]$. The approximation accuracy of decision class D_j $(j = 1, 2, \ldots, t)$ with respect to $\sum_{i=1}^{S} A_i$ under different information levels in the WGM-IVDTRS model has the following conclusions:*

$acc^{\varphi_1}(D_j) \ge acc^{\varphi_2}(D_j)$.

Proof. When $\varphi_1 \le \varphi_2$, the lower approximation $\underline{WGM}^{\varphi_1}_{\sum_{i=1}^{s} A_i}(D_j)$ under information level φ_1 contains the lower approximation $\underline{WGM}^{\varphi_2}_{\sum_{i=1}^{s} A_i}(D_j)$ under φ_2, and the upper approximation $\overline{WGM}^{\varphi_1}_{\sum_{i=1}^{s} A_i}(D_j)$ under φ_1 is contained in the upper approximation $\overline{WGM}^{\varphi_2}_{\sum_{i=1}^{s} A_i}(D_j)$ under φ_2. Therefore, the conclusion is clearly established.

By analyzing the three-way decision regions, we obtained weighted generalized, optimistic and pessimistic multi-granulation decision rules.

Rule 4.1 In the WGM-IVDTRS model, for any $X \subseteq U$, weighted generalized multi-granulation decision rules are listed as follows:

(P_G) If $\sum\{\varpi_i | P(X|[x]_{A_i}) \geq \alpha_i\} \geq \varphi$, then decide $x \in pos(X)$;

(B_G) If $\sum\{\varpi_i | P(X|[x]_{A_i}) > \beta_i\} > 1-\varphi$ and $\sum\{\varpi_i | P(X|[x]_{A_i}) \geq \alpha_i\} < \varphi$, then decide $x \in bnd(X)$;

(N_G) If $\sum\{\varpi_i | P(X|[x]_{A_i}) > \beta_i\} \leq 1-\varphi$, then decide $x \in neg(X)$.

According to rule 4.1, if the sum of the weights of granulations satisfying $P(X|[x]_{A_i}) \geq \alpha_i$ is not smaller than φ, then decide $x \in pos(X)$; if the sum of the weights of granulations satisfying $P(X|[x]_{A_i}) \geq \alpha_i$ is smaller than φ and the sum of the weights of granulations satisfying $P(X|[x]_{A_i}) > \beta_i$ is greater than $1-\varphi$, then decide $x \in bnd(X)$; if the sum of the weights of granulations satisfying $P(X|[x]_{A_i}) > \beta_i$ is not greater than $1-\varphi$, then decide $x \in neg(X)$.

Rule 4.2 In the WOM-IVDTRS and WPM-IVDTRS models, for any $X \subseteq U$, weighted optimistic and pessimistic multi-granulation decision rules are listed as follows:

(P_O) If $|A_i : P(X|[x]_{A_i}) \geq \alpha_i| \geq 1$, then decide $x \in pos(X)$;

(B_O) If $|A_i : P(X|[x]_{A_i}) > \beta_i| = s$ and $|A_i : P(X|[x]_{A_i}) \geq \alpha_i| = 0$, then decide $x \in bnd(X)$;

(N_O) If $|A_i : P(X|[x]_{A_i}) > \beta_i| < s$, then decide $x \in neg(X)$;

(P_P) If $|A_i : P(X|[x]_{A_i}) \geq \alpha_i| = s$, then decide $x \in pos(X)$;

(B_P) If $|A_i : P(X|[x]_{A_i}) > \beta_i| \geq 1$ and $|A_i : P(X|[x]_{A_i}) \geq \alpha_i| < s$, then decide $x \in bnd(X)$;

(N_P) If $|A_i : P(X|[x]_{A_i}) > \beta_i| = 0$, then decide $x \in neg(X)$.

Under the optimistic circumstance, if at least one granulation satisfies $P(X|[x]_{A_i}) \geq \alpha_i$, then decide $x \in pos(X)$; if all granulations satisfy $P(X|[x]_{A_i}) > \beta_i$, and all granulations do not satisfy $P(X|[x]_{A_i}) \geq \alpha_i$, then decide $x \in bnd(X)$; if not all granulations satisfy $P(X|[x]_{A_i}) > \beta_i$, then decide $x \in neg(X)$. And under the pessimistic circumstance, if all granular structures satisfy $P(X|[x]_{A_i}) \geq \alpha_i$, then decide $x \in pos(X)$; if at least one granular structure satisfies $P(X|[x]_{A_i}) > \beta_i$ and not all granular structures satisfy $P(X|[x]_{A_i}) \geq \alpha_i$, then decide $x \in bnd(X)$; if all granular structures do not satisfy $P(X|[x]_{A_i}) > \beta_i$, then decide $x \in neg(X)$.

In the following, the relationships among the proposed WGM-IVDTRS model and other rough set models are explored.

• From the perspective of granulation importance, the WGM-IVDTRS model is a generalization of generalized multi-granulation rough sets (GMRS), optimistic multi-granulation rough sets (OMRS) and pessimistic multi-granulation rough sets (PMRS).

• From the perspective of granulation selection, the WGM-IVDTRS model is a generalization of the OMRS and PMRS models.

• From the perspective of multiple granulations, the WGM-IVDTRS model is a generalization of the IVDTRS model.

• From the perspective of loss functions, the WGM-IVDTRS model is a generalization of the decision-theoretic rough set (DTRS) model.

4.2. The granulation weighted method based on classification accuracy

It is well known that all granulations are not equally important in practical problems. The information gain produced by different granulations is different. Certainly, the support degree of different granulations about relevant decisions is also different. The existing granulation weighted methods are mainly based on the upper and lower approximations characterized by equivalence classes under different granulations. In the approximation space with the universe U and attribute set AT, their computational complexities are all $O(|AT| * |U|^2)$, which limits their applications in large data sets. Moreover, these methods consider all the attributes contained in the granulation at the same time when weighting each granulation, and they lack generalization ability for small changed data. Rough set theory and decision tree are all learning methods based on labeled data. Decision tree focuses on finding the most useful attribute information to distinguish objects of different classes, and the time complexity $O(|AT| * |U|)$ is relatively low. After many cross validation, decision tree methods can respond more objectively and flexibly to the importance of different granulations. Therefore, a granulation weighted method is proposed based on the classification accuracy of decision tree learning. Considering the flexibility and objectivity of data analysis, in this paper we analyze and classify data by the improved decision tree algorithm ID_3 which selects the optimal partition attribute by information gain ratio. In the classification process, we use 70% of the data set as the training set S, and the rest as the test set T. The decision tree of training set S and the classification accuracy ρ_i of decision tree on test set T can be obtained under granulation A_i $(i = 1, 2, \ldots, s)$.

Definition 4.4. Let $I = (U, AT \cup DT, V, F)$ be an information system, $A_i \subseteq AT$, $i = 1, 2, \ldots, s$ $(s \leq 2^{AT})$, ρ_i is the classification accuracy under granulation A_i $(i = 1, 2, \ldots, s)$. The weight ϖ_i of granulation A_i $(i = 1, 2, \ldots, s)$ is defined as

$$\varpi_i = \rho_i / \sum_{i=1}^{s} \rho_i. \tag{19}$$

where ρ_i is the classification accuracy obtained from decision tree learning under granulation A_i.

In order to demonstrate the proposed granulation weighted method more intuitively, we illustrate the calculation process of this weighted method through a practical example.

Example 2. Let $I = (U, AT \cup DT, V, F)$ be the watermelon data set 2.0 derived from Zhou Zhihua's machine learning [50], detailed information is shown in Table 4. The universe is $U = \{x_1, x_2, \ldots, x_{17}\}$, the conditional attribute set is $AT = \{a_1, a_2, a_3, a_4, a_5, a_6\} = \{color, root, stroke, texture, navel, touch\}$ and the decision attribute set is $DT = \{d\} = \{good\ melon\}$.

First of all, we describe how to calculate information gain. In an information system $I = (U, AT \cup DT, V, F)$, let $U/DT = \{D_1, D_2, \ldots, D_t\}$, $U/a_i = \{D_1^{a_i}, D_2^{a_i}, \ldots, D_h^{a_i}\}$, and $D_k/a_i = \{D_{k1}^{a_i}, D_{k2}^{a_i}, \ldots, D_{kn}^{a_i}\}$, where h denotes the number of attribute values of all objects in U under attribute a_i ($i = 1, 2, \ldots, |AT|$), n denotes the number of attribute values of all objects in D_k under attribute a_i $(k = 1, 2, \ldots, t)$. The information entropy of the universe U about decision DT [50] is defined as $Ent(U, DT) = -\sum_{k=1}^{t} \frac{|D_k|}{|U|} log_2 \frac{|D_k|}{|U|}$. The information gain obtained by the partition of U under attribute a_i $(a_i \in AT)$ is defined as $Gain(U, a_i) = Ent(U, DT) - \sum_{j=1}^{h} \frac{|D_j^{a_i}|}{|U|} Ent(D_j^{a_i}, DT)$. The information gain ratio of attribute a_i is defined as $Gain\text{-}ratio(U, a_i) = \frac{Gain(U, a_i)}{Iv(a_i)}$, where $Iv(a_i) = -\sum_{j=1}^{h} \frac{|D_j^{a_i}|}{|U|} log_2 \frac{|D_j^{a_i}|}{|U|}$ is the intrinsic value of attribute a_i in the universe U. According to the information in Table 4, we have $U/d = \{D_1, D_2\} = \{\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}, \{x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}\}\}$. Then the information entropy of the universe U about decision d is $Ent(U, d) = -(\frac{8}{17} log_2 \frac{8}{17} + \frac{9}{17} log_2 \frac{9}{17}) = 0.9975$. According to $U/a_1 = \{D_1^{a_1}, D_2^{a_1}, D_3^{a_1}\} = \{\{x_1, x_4, x_6, x_{10}, x_{13}, x_{17}\}, \{x_2, x_3, x_7, x_8, x_9, x_{15}\}, \{x_5, x_{11}, x_{12}, x_{14}, x_{16}\}\}$, $D_1^{a_1}/d = \{\{x_1, x_4, x_6\}, \{x_{10}, x_{13}, x_{17}\}\}$, $D_2^{a_1}/d = \{\{x_2, x_3, x_7, x_8\}, \{x_9, x_{15}\}\}$ and $D_3^{a_1}/d = \{\{x_5\}, \{x_{11}, x_{12}, x_{14}, x_{16}\}\}$, the information entropy of $D_1^{a_1}, D_2^{a_1}, D_3^{a_1}$ about decision d is $Ent(D_1^{a_1}, d) = -(\frac{3}{6} log_2 \frac{3}{6} + \frac{3}{6} log_2 \frac{3}{6}) = 1$, $Ent(D_2^{a_1}, d) = -(\frac{4}{6} log_2 \frac{4}{6} + \frac{2}{6} log_2 \frac{2}{6}) = 0.9183$ and $Ent(D_3^{a_1}, d) = -(\frac{1}{5} log_2 \frac{1}{5} + \frac{4}{5} log_2 \frac{4}{5}) = 0.7219$. Therefore, the information gain of attribute a_1 is $Gain(U, a_1) = 0.9975 - (\frac{6}{17} Ent(D_1^{a_1}, d) +$

Table 4
A decision system I.

U	color	root	stroke	texture	navel	touch	good melon
x_1	turquois [a]	curled up[a]	muffled[a]	clear [a]	concave[a]	hard and smooth[a]	good[a]
x_2	dark [a]	curled up	dull[a]	clear	concave	hard and smooth	good
x_3	dark	curled up	muffled	clear	concave	hard and smooth	good
x_4	turquois	curled up	dull	clear	concave	hard and smooth	good
x_5	pale[a]	curled up	muffled	clear	concave	hard and smooth	good
x_6	turquois	slightly curled up[a]	muffled	clear	slightly dented[a]	soft and sticky[a]	good
x_7	dark	slightly curled up	muffled	slightly blurry[a]	slightly dented	soft and sticky	good
x_8	dark	slightly curled up	muffled	clear	slightly dented	hard and smooth	good
x_9	dark	slightly curled up	dull	slightly blurry	slightly dented	hard and smooth	bad[a]
x_{10}	turquois	stiff [a]	crisp[a]	clear	flat[a]	soft and sticky	bad
x_{11}	pale	stiff	crisp	fuzzy[a]	flat	hard and smooth	bad
x_{12}	pale	curled up	muffled	fuzzy	flat	soft and sticky	bad
x_{13}	turquois	slightly curled up	muffled	slightly blurry	concave	hard and smooth	bad
x_{14}	pale	slightly curled up	dull	slightly blurry	concave	hard and smooth	bad
x_{15}	dark	slightly curled up	muffled	clear	slightly dented	soft and sticky	bad
x_{16}	pale	curled up	muffled	fuzzy	flat	hard and smooth	bad
x_{17}	turquois	curledup	dull	slightly blurry	slightly dented	hard and smooth	bad

[a]Facilitate the description, we abbreviate *turquois*, *dark* and *pale* describing color as *TU*, *DA*, *PA*; *curled up*, *slightly curled up*, *stiff* describing root as *CU*, *SU*, *ST*; *muffled*, *dull*, *crisp* describing stroke as *MU*, *DU*, *CR*; *clear*, *slightly blurry*, *fuzzy* describing texture as *CL*, *SB*, *FU*; *concave*, *slightly dented*, *flat* describing navel as *CO*, *SD*, *FL*; *hard and smooth*, *soft and sticky* describing touch as *HS*, *SS*; *good*, *bad* as *G* and *B*.

$\frac{6}{17}Ent(D_2^{a_1}, d) + \frac{5}{17}Ent(D_1^{a_1}, d)) = 0.1081$. The intrinsic value of a_1 is $IV(a_1) = -(\frac{6}{17}log_2\frac{6}{17} + \frac{6}{17}log_2\frac{6}{17} + \frac{5}{17}log_2\frac{5}{17}) = 1.5799$. Then the information gain ratio of attribute a_1 is $Gain\text{-}ratio(U, a_1) = 0.0684$. Similarly, information gain ratios of other attributes are obtained, namely $Gain\text{-}ratio(U, a_2) = 0.1018$, $Gain\text{-}ratio(U, a_3) = 0.1056$, $Gain\text{-}ratio(U, a_4) = 0.2631$, $Gain\text{-}ratio(U, a_5) = 0.1867$ and $Gain\text{-}ratio(U, a_6) = 0.0069$.

Secondly, we introduce how to calculate the weights of different granulations. Using the hold-out method, the universe is randomly divided into two parts, where 70% of objects form a training set S_i, and the other objects as the corresponding test set T_i. In order to make the estimated results stable and reliable, we conducted many random partitions and experiments, then take the average of multiple classification accuracy under the same granulation as the weight of this granulation. Taking into account the scale of the case data set, we carried out four experiments. Four granulations are randomly selected, which are $A_1 = \{a_1, a_3, a_4, a_5\}, A_2 = \{a_1, a_2, a_4, a_6\}, A_3 = \{a_1, a_3, a_6\}$ and $A_4 = \{a_2, a_3, a_4, a_5, a_6\}$. The training sets and test sets of these four experiments are shown in Table 5.

Decision rules of different granulations may be different by decision tree learning on the same training set. In the four experiments, decision rules learned by the improved decision tree algorithm are shown in Figs. 1–4, respectively. In Figs. 1–4, the contents of yellow rectangles denote the attributes selected according to the information gain ratio, the letters on the blue flow line describe the characteristics of the attributes adjacent to them, and the contents in the red circles denote the final decisions. The red small circles are called the leaf nodes of decision tree. Decision tree learning is a top-down learning process. By observing, the total number of decision rules can be obtained based on the number of leaf nodes in each decision tree and specific decision rules can also be obtained by following the path from the root attribute to the leaf node.

For easier understanding, we describe the decision tree in Fig. 1. For example, in the subgraph (a) of Fig. 1, it is apparent that the maximum information gain ratio under the granulation A_1 is **navel**. If the **navel** of the watermelon is **concave**, then the **texture** of the watermelon needs further observation; if the **navel** of the watermelon is **slightly dented** or **flat**, then this watermelon is a **bad melon**. If the **navel** of the watermelon is **concave** and the **texture** of the watermelon is **clear** or **fuzzy**, then this watermelon is a **good melon**. If the **navel** of the watermelon is **concave** and the **texture** of the watermelon is **slightly blurry**, then this watermelon is a **bad melon**. There are **5** decision rules learned under granulation A_1. The total number of decision rules under the remaining three granulations A_2, A_3, A_4 is **11,6,9**, respectively.

Finally, the results of four experiments on test sets are shown in Table 5, where Gr, EO, CA are abbreviations for different granulations, classification error objects and classification accuracy. According to the classification accuracy of the four tests, the average classification accuracy of each granulation can be obtained, namely, $\rho_{A_1} = \frac{17}{28}$, $\rho_{A_2} = \frac{4}{7}$, $\rho_{A_3} = \frac{3}{7}$, $\rho_{A_4} = \frac{6}{7}$. Therefore, according to Definition 4.4, the weights of four granulations are $\varpi_1 = \frac{17}{69}, \varpi_2 = \frac{16}{69}, \varpi_3 = \frac{12}{69}, \varpi_4 = \frac{24}{69}$.

4.3. Three types of the WGM-IVDTRS model

Considering that the final goal of interval-valued decision-theoretic rough sets is to find the minimum expected loss of each object and it is very time-consuming to make intervals comparison under multiple granulations, we propose three types of the WGM-IVDTRS model based on different methods of determining threshold parameters. The first two methods transform the interval-valued loss function into the real values by using different transformation functions, then obtain directly parameters; the third approach is to find the optimal threshold parameters that make the information uncertainty minimum. Details are described in the following.

Firstly, considering that the certain ranking method can reflect the risk preferences of different decision makers, we propose weighted generalized multi-granulation certain ranking interval-valued decision-theoretic rough sets (**WGM-IVDTRSC**). The upper and lower approximations of concept X with respect to $\sum_{i=1}^{s} A_i$ in the WGM-IVDTRSC model are defined as

$$\overline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | \sum_{i=1}^{s} WUS_X^{A_i}(x) > 1 - \varphi\},$$

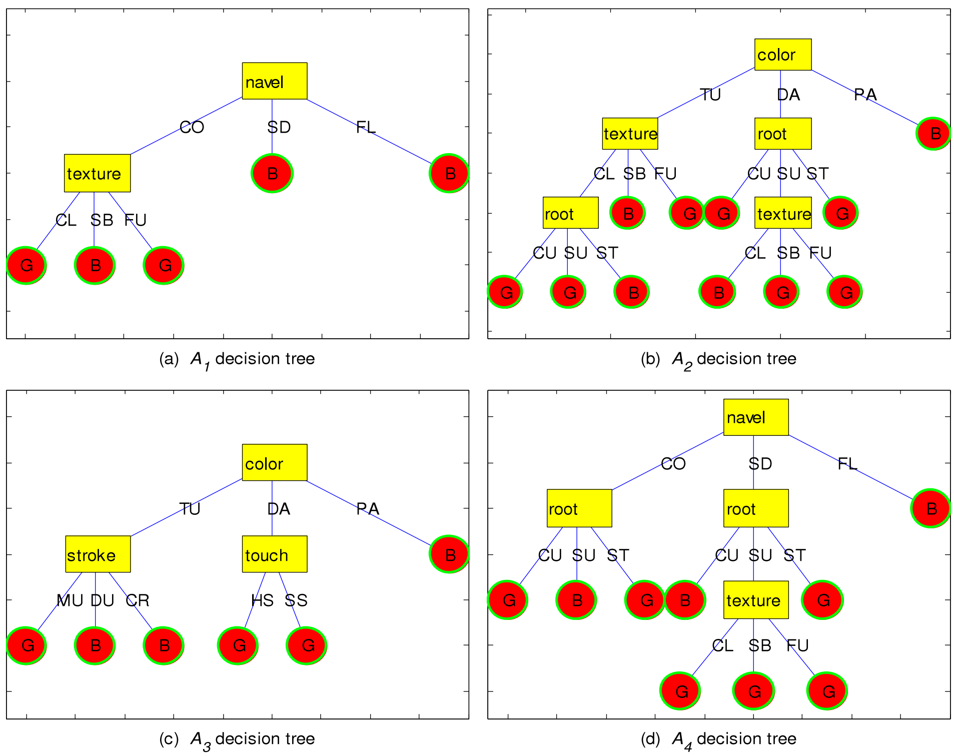
$$\underline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | \sum_{i=1}^{s} WLS_X^{A_i}(x) \geq \varphi\},$$

where $WUS_X^{A_i}(x)$ and $WLS_X^{A_i}(x)$ are the weighted upper and lower support characteristic functions of x with respect to X under granulation A_i, and α_C, β_C in the weighted support characteristic functions are calculated by formula (2).

Secondly, considering the good performance of geometry average interval sorting method, we propose weighted generalized multi-granulation geometry average sorting interval-valued decision-theoretic rough sets (**WGM-IVDTRSG**). For any subset X

Table 5
The information about four experiments.

1st	S_1 T_1	$x_1, x_2, x_3, x_4, x_5, x_9, x_{10}, x_{11}, x_{12}, x_{13}$ $x_6, x_7, x_8, x_{14}, x_{15}, x_{16}, x_{17}$	2nd	S_2 T_2	$x_2, x_3, x_4, x_5, x_6, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}$ $x_1, x_7, x_8, x_9, x_{15}, x_{16}, x_{17}$
Gr	*EO*	*CA*	*Gr*	*EO*	*CA*
A_1	x_6, x_7, x_8	4/7	A_1	x_7, x_{15}	5/7
A_2	x_7, x_8, x_{14}	4/7	A_2	x_7, x_8, x_9	4/7
A_3	x_6, x_7, x_{14}, x_{17}	3/7	A_3	x_1, x_7, x_9, x_{17}	3/7
A_4	x_{14}	6/7	A_4	x_9	6/7
3rd	S_3 T_3	$x_3, x_4, x_5, x_6, x_7, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}$ $x_1, x_2, x_8, x_9, x_{10}, x_{11}, x_{17}$	4th	S_4 T_4	$x_4, x_5, x_6, x_7, x_8, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}$ $x_1, x_2, x_3, x_9, x_{10}, x_{11}, x_{12}$
Gr	*EO*	*CA*	*Gr*	*EO*	*CA*
A_1	x_8, x_9, x_{10}, x_{17}	3/7	A_1	x_9, x_{10}	5/7
A_2	x_2, x_8, x_9	4/7	A_2	x_2, x_3, x_9	4/7
A_3	x_1, x_2, x_9, x_{17}	3/7	A_3	x_1, x_2, x_9, x_{12}	3/7
A_4	x_9	6/7	A_4	x_9	6/7



(a) A_1 decision tree (b) A_2 decision tree (c) A_3 decision tree (d) A_4 decision tree

Fig. 1. Decision rules learned from the first experiment.

of U, the upper and lower approximations of X with respect to $\sum_{i=1}^{s} A_i$ in the WGM-IVDTRSG model are defined as

$$\overline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | \sum_{i=1}^{s} WUS_X^{A_i}(x) > 1 - \varphi\},$$

$$\underline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | \sum_{i=1}^{s} WLS_X^{A_i}(x) \geq \varphi\},$$

where parameters α_G, β_G in the weighted upper and lower support characteristic functions are calculated by

$$\alpha_G = \frac{G_\eta(\widetilde{\lambda}_{PN}) - G_\eta(\widetilde{\lambda}_{BN})}{G_\eta(\widetilde{\lambda}_{PN}) - G_\eta(\widetilde{\lambda}_{BN}) + G_\eta(\widetilde{\lambda}_{BP}) - G_\eta(\widetilde{\lambda}_{PP})},$$
$$\beta_G = \frac{G_\eta(\widetilde{\lambda}_{BN}) - G_\eta(\widetilde{\lambda}_{NN})}{G_\eta(\widetilde{\lambda}_{BN}) - G_\eta(\widetilde{\lambda}_{NN}) + G_\eta(\widetilde{\lambda}_{NP}) - G_\eta(\widetilde{\lambda}_{BP})}. \quad (20)$$

Symbols $G_\eta(\widetilde{\lambda}_{PP})$, $G_\eta(\widetilde{\lambda}_{PN})$, $G_\eta(\widetilde{\lambda}_{BP})$, $G_\eta(\widetilde{\lambda}_{BN})$, $G_\eta(\widetilde{\lambda}_{NP})$ and $G_\eta(\widetilde{\lambda}_{NN})$ are geometry average sorting functions of intervals $\widetilde{\lambda}_{PP}$, $\widetilde{\lambda}_{PN}$, $\widetilde{\lambda}_{BP}$, $\widetilde{\lambda}_{BN}$, $\widetilde{\lambda}_{NP}$ and $\widetilde{\lambda}_{NN}$.

Finally, considering the objectivity of the data response information, and from the machine learning point of view, we propose weighted generalized multi-granulation optimization interval-valued decision-theoretic rough sets (**WGM-IVDTRSO**). The upper and lower approximations of X with respect to $\sum_{i=1}^{s} A_i$ in the WGM-IVDTRSO model are defined as

$$\overline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | \sum_{i=1}^{s} WUS_X^{A_i}(x) > 1 - \varphi\},$$

$$\underline{WGM}_{\sum_{i=1}^{s} A_i}(X) = \{x \in U | \sum_{i=1}^{s} WLS_X^{A_i}(x) \geq \varphi\},$$

where parameters α_O, β_O in the weighted upper and lower support characteristic functions make the overall uncertainty of three rough regions $H(\pi_{DT} | \pi_{(\alpha_O, \beta_O)})$ minimum, which can be calculated by formulas (3) and (4).

In order to deeply understand the proposed WGM-IVDTRSC, WGM-IVDTRSG, WGM-IVDTRSO theory, we introduce an example to explain the process of solving the approximations of these three models.

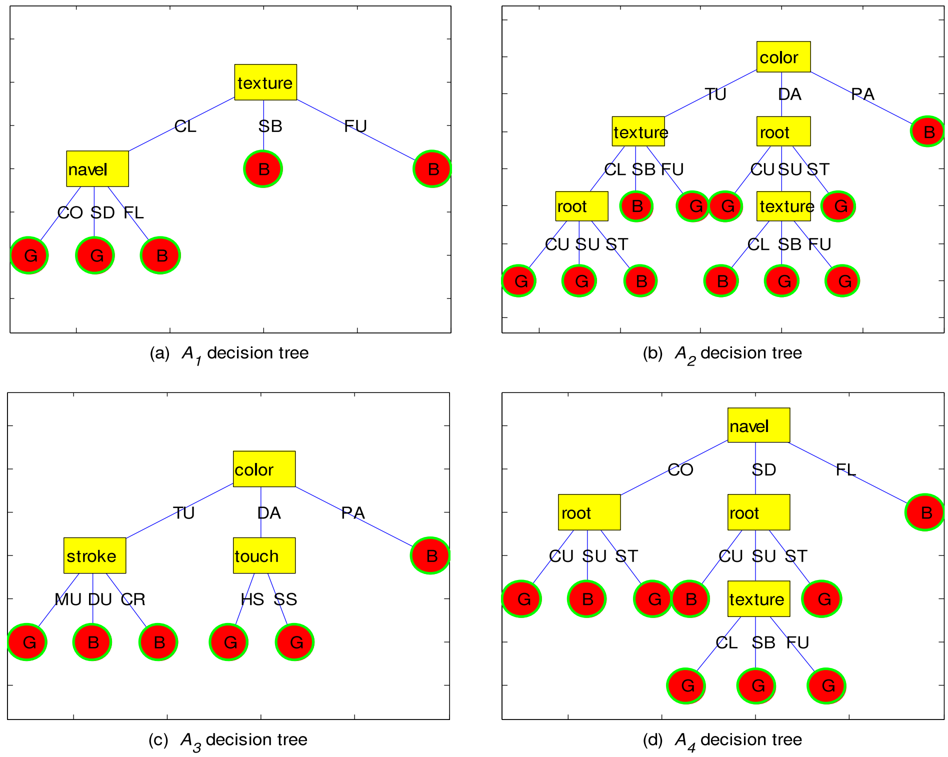


Fig. 2. Decision rules learned from the second experiment.

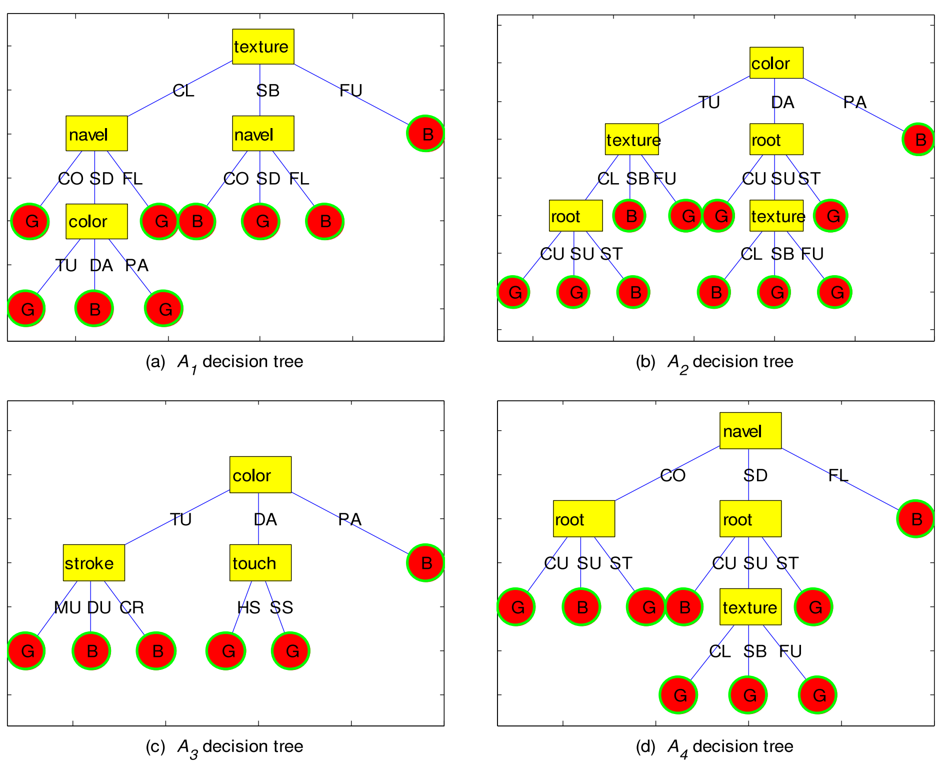
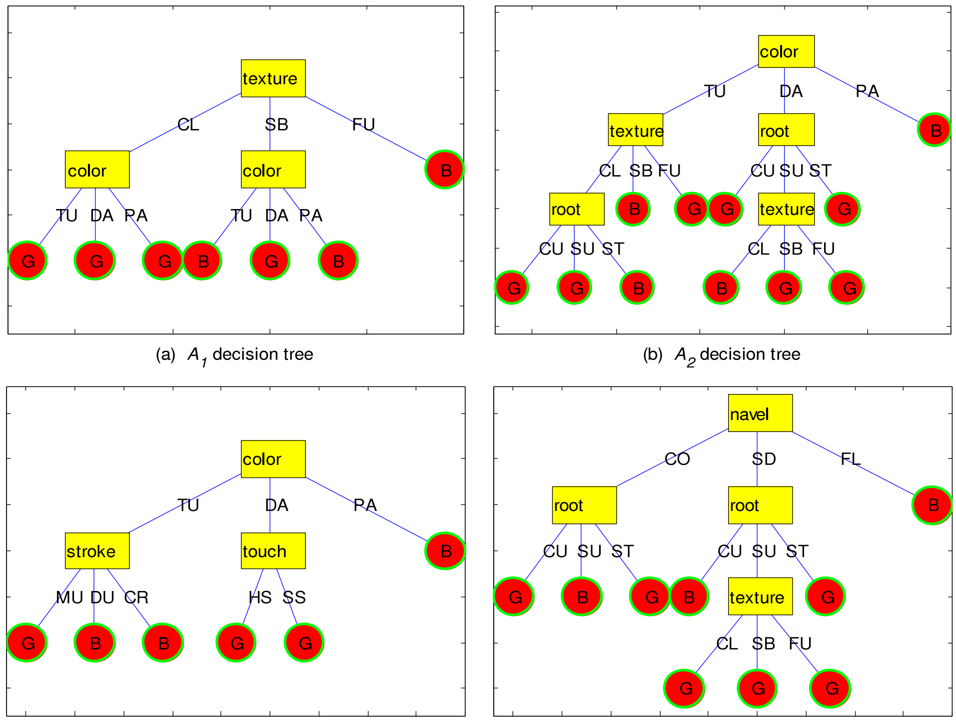


Fig. 3. Decision rules learned from the third experiment.

Example 3. Following Example 2, we calculate the upper and lower approximations of the proposed WGM-IVDTRSC, WGM-IVDTRSG, WGM-IVDTRSO models. The values of parameters in interval-valued loss function are the same as that in Example 1. The randomly selected decision class X is $\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ and the information level φ is 0.8. It is important to point out that the weights of four granulations are $\varpi_1 = \frac{17}{69}, \varpi_2 = \frac{16}{69}, \varpi_3 = \frac{12}{69}, \varpi_4 = \frac{24}{69}$.

In the WGM-IVDTRSC model, the risk preference parameter θ is set to 0.6. According to formula (2), threshold parameters are



(a) A_1 decision tree (b) A_2 decision tree (c) A_3 decision tree (d) A_4 decision tree

Fig. 4. Decision rules learned from the fourth experiment.

$\alpha_C = 0.7938$, $\beta_C = 0.4316$. The upper and lower approximations of X with respect to A_i $(i = 1, 2, 3, 4)$ are
$\overline{A}_1(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_{15}\}$,
$\underline{A}_1(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$;
$\overline{A}_2(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$,
$\underline{A}_2(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$;
$\overline{A}_3(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{13}, x_{15}, x_{16}, x_{17}\}$,
$\underline{A}_3(X) = \{x_3, x_6, x_8\}$;
$\overline{A}_4(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_{15}\}$,
$\underline{A}_4(X) = \{x_1, x_2, x_3, x_4, x_5, x_7\}$.

Therefore, the upper and lower approximations of X with respect to $\sum_{i=1}^{4} A_i$ in WGM-IVDTRSC are $\overline{WGM}_{\sum_{i=1}^{4} A_i}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_{15}\}$, $\underline{WGM}_{\sum_{i=1}^{4} A_i}(X) = \{x_1, x_2, x_3, x_4, x_5, x_7\}$.

Then the classification error rate of X can by calculated by formula (18), namely $e_C = \frac{|X\cap neg(X)|+|\sim X\cap pos(X)|}{|U|} = 0$.

In the WGM-IVDTRSG model, the credibility parameter η is set to 0.6. By formula (20), we have $\alpha_G = 0.7609$, $\beta_G = 0.5055$. Then the upper and lower approximations of X with respect to A_i $(i = 1, 2, 3, 4)$ are
$\overline{A}_1(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$,
$\underline{A}_1(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$;
$\overline{A}_2(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$,
$\underline{A}_2(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$;
$\overline{A}_3(X) = \{x_3, x_6, x_8\}$,
$\underline{A}_3(X) = \{x_3, x_6, x_8\}$;
$\overline{A}_4(X) = \{x_1, x_2, x_3, x_4, x_5, x_7, x_8\}$,
$\underline{A}_4(X) = \{x_1, x_2, x_3, x_4, x_5, x_7, x_8\}$.

Therefore, the upper and lower approximations of X with respect to $\sum_{i=1}^{4} A_i$ in WGM-IVDTRSG are $\overline{WGM}_{\sum_{i=1}^{4} A_i}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$, $\underline{WGM}_{\sum_{i=1}^{4} A_i}(X) = \{x_1, x_2, x_3, x_4, x_5, x_7\}$. Its classification error rate is $e_G = 0$.

In the WGM-IVDTRSO model, the optimal threshold parameters are found by particle swarm optimization algorithm (PSO). The parameters of PSO are as follows: the learning factors c_1, c_2 are $c_1 = c_2 = 2$, the fitness function is $H(\pi_{DT}|\pi_{\alpha_O,\beta_O})$, the number of particles is $Num = 20$, the maximum number of generations as a generation stopping criterion is $Max = 200$. The dimension of every particle is $t = 6$, namely $\lambda_{PP}, \lambda_{PN}, \lambda_{BP}, \lambda_{BN}, \lambda_{NP}, \lambda_{NN}$. The position of the particle is $x(i,j)$ and its corresponding velocity is $\upsilon(i,j)$ $(i = 1, 2, \ldots, Num; j = 1, 2, \ldots, t)$. The update of the velocity is realized in the form of $\upsilon(i,j) = \omega(p) * \upsilon(i,j) + c_1 * rand * (y(i,j) - x(i,j)) + c_2 * rand * (pg(j) - x(i,j))$, where $\omega(p) = 0.9 - (0.9 - 0.4) * p/Max$ is dynamic inertia weight, p is iterations, $y(i,j)$ is the position of local best solution and $pg(j)$ is the position of global best solution. The next position of the particle is computed as $x(i,j) = x(i,j) + \upsilon(i,j)$. The optimal result of $H(\pi_{DT}|\pi_{\alpha_O,\ \beta_O})$ is 0.1176, in which $\lambda_{PP} = 2.0215$, $\lambda_{PN} = 4.7642$, $\lambda_{BP} = 2.9568$, $\lambda_{BN} = 2.4672$, $\lambda_{NP} = 4.0254$, $\lambda_{NN} = 1.5475$. The optimal thresholds are $\alpha_{opt} = 0.7105$ and $\beta_{opt} = 0.4625$.

Then the upper and lower approximations of X with respect to A_i $(i = 1, 2, 3, 4)$ are
$\overline{A}_1(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7 x_8, x_{15}\}$,
$\underline{A}_1(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$;
$\overline{A}_2(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$,
$\underline{A}_2(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$;
$\overline{A}_3(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7 x_8, x_9, x_{13}, x_{15}, x_{16}, x_{17}\}$,
$\underline{A}_3(X) = \{x_3, x_6, x_8\}$;
$\overline{A}_4(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_{15}\}$,
$\underline{A}_4(X) = \{x_1, x_2, x_3, x_4, x_5, x_7, x_8\}$.

Therefore, the upper and lower approximations of X with respect to $\sum_{i=1}^{4} A_i$ in the WGM-IVDTRSO model are $\overline{WGM}_{\sum_{i=1}^{4} A_i}(X)$

$= \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_{15}\}$, and $\underline{WGM}_{\sum_{i=1}^{4} A_i}(X) = \{x_1, x_2, x_3, x_4, x_5, x_7\}$. The classification error rate of X in the WGM-IVDTRSO model is $e_O = 0$.

5. Experimental results and analyses

There are three different weighted generalized multi-granulation interval-valued decision-theoretic rough set models proposed, namely weighted generalized multi-granulation certain ranking interval-valued decision-theoretic rough sets (WGM-IVDTRSC), weighted generalized multi-granulation geometry average sorting interval-valued decision-theoretic rough sets (WGM-IVDTRSG) and weighted generalized multi-granulation optimization interval-valued decision-theoretic rough sets (WGM-IVDTRSO). First, the necessity, feasibility and effectiveness of the weighted method based on the classification accuracy of decision tree learning are verified by comparing it with multi-granulation decision-theoretic methods under the interval-valued loss function and WGM-IVDTRS models based on different granulation weighted methods. Then the proposed models based on the classification accuracy of decision tree learning (WGM-IVDTRSC, WGM-IVDTRSG, WGM-IVDTRSO) are compared by experimental analyses and some criteria for choosing a suitable analysis method in the interval-valued loss environment are obtained. The classification error rate is introduced to evaluate performances of these decision models. The smaller of the classification error rate implies that the performance is better.

Nine data sets are downloaded from UCI Machine Learning Repository (http://archive.ics.uci.edu/ml/datasets.html). Experiments are performed on a computer with 2.6 GHz CPU, 8.0 GB of memory and 64-bit Windows 10, and have been implemented through MATLAB 2015B. Detailed data information is shown in Table 6.

5.1. Pretreatment of the data sets and analyses of the WGM-IVDTRS models

As the IVDTRS model focuses on the dichotomy problem, so some decision classes need to be merged. In addition, different probability distributions and inconsistencies of data information will affect the performance of the WGM-IVDTRS models. In order to exert the advantage of three-way decisions [6], we delete some attributes of data sets similar to the preprocessing method in literature [16]. The preprocessing method is: firstly, the condition attributes are deleted randomly in the data sets; secondly, if the decision classes are more than two in these data sets, then some decision classes are merged and the concepts are defined. The detailed results are shown in Table 7. And the distribution of data after preprocessing is shown in Table 8.

There are three parameters affecting the performance of the WGM-IVDTRS model, namely parameters α, β, information level φ and granulation weight ω. (1) The parameters α, β mainly depend on the interval-valued loss function. The certain ranking and optimization methods proposed by Liang [16] and geometry average sorting method can be used to determine the parameters α, β. In the certain ranking and geometry average sorting methods, the parameters α, β mainly depend on the risk attitudes of decision makers, namely parameter θ, η. Moreover, as the degree of optimism of risk attitude decreases (namely θ, η increase), the parameters α increases and β decreases [16]. This paper is mainly concerned with moderate risk attitude. So we set $\theta = 0.5$ and $\eta = 0.6$. Of course, other values can be set according to different risk preferences. Based on the above three parameter determination methods, three pairs of decision risk parameters can be obtained. (2) For the information level parameter φ, it is true that with the increase of information level, the upper approximation of a concept becomes larger, and the lower approximation becomes smaller according to proposition 4.3. That is to say, the positive and negative regions of the concept become smaller. Therefore, the classification error rate of the concept becomes smaller with the increase of information level. In this section, we set φ to a value between 0.6 and 1 in steps of 0.2. (3) For granulation weights, the importance, feasibility and effectiveness of the weighted method based on the classification accuracy of decision tree learning are explored by comparing it with three non-weighted multi-granulation decision making models and three granulation weighted methods. For easy descriptions, under the interval-valued loss function, non-weighted multi-granulation decision models [18] proposed by Qian et al. have become corresponding mean, optimistic and pessimistic multi-granulation interval-valued decision-theoretic rough sets, which can be abbreviated as MM-IVDTRS, OM-IVDTRS and PM-IVDTRS, respectively. In the absence of ambiguity, in the following tables we abbreviate them as MM, OM and PM, respectively. At the same time, the classification error rates of the WGM-IVDTRS model under different weighted methods are compared. In addition to the classification accuracy weighted method mentioned in this paper, the other weighted methods we choose are approximation accuracy weighted method, approximation quality weighted method and granulation entropy weighted method [31], respectively. Similarly, approximation accuracy weighted, approximation quality weighted and granulation entropy weighted methods are abbreviated as AAW, AQW and GEW, respectively.

In the following experiments, four granulations are selected in each data set. In order to verify the validity of the WGM-IVDTRS model, we sequentially select granulations with the almost same number of condition attributes in each data set. Under the finest partition (namely for each object, its equivalence class contains only itself), the probability information of the concept is either 0 or 1. To be more general, we try to avoid this extreme situation that the partitions under these granulations are the finest partition. The results of granulation selection are shown in Table 9. For each data set, the loss values are generated randomly in the interval $[0, 1]$ according to the basic constraints of the loss function. The generation method is similar to literature [16]. The information of loss functions is shown in Table 10.

Then the decision risk parameters are calculated according to certain ranking, geometry average sorting and optimization methods. The certain ranking and geometry average sorting methods are mainly based on the transformation of interval function to calculate the parameters. The optimization method is based on particle swarm optimization to find the parameters that minimize the overall uncertainty of the three regions. The specific results are shown in Table 11.

5.2. Comparison of three WGM-IVDTRS models and other models

In order to evaluate performances of three WGM-IVDTRS models and find the criteria for selecting a suitable analysis method under different practical environments, the classification error rates of concepts are compared in the WGM-IVDTRS models, the MM, OM, PM models and the AAW, AQW, GEW methods using the above nine data sets.

Firstly, the weighted method based on the classification accuracy of decision tree learning is introduced in detail. The weights of four granulations are calculated on each data set. Each data set is randomly divided into four parts and four tests are performed on each data set. In four tests, we take one out of four parts as the test set in turn and the remaining three parts as the training set. The average classification accuracy of the four tests is then calculated as the classification accuracy of each granulation. Finally, based on the classification accuracy, the weight

Table 6
Data descriptions.

Data sets	Abbreviation	Samples	Condition attributes	Decision classes
Tic-Tac-Toe Endgame	Tic.	958	9	2
Contraceptive Method Choice	Cmc.	1473	9	3
Car Evaluation	Car.	1728	6	4
Chess(King-Rook vs. King-Pawn)	Che.	3196	36	2
Mushroom	Mus.	8124	22	2
Nursery	Nur.	12960	8	5
Default of Credit Card Clients	Def.	30000	23	2
Bank Marketing	Ban.	45211	16	2
Connect-4	Con.	67557	42	3

Table 7
The pretreatment of nine data sets.

Data	Deleted attributes	Attributes	Merged classes	Final classes	Concept C
Tic.	Bottom-left-square, bottom-middle-square, bottom-right-square	6		{positive}, {negative}	{positive}
Cmc.	Wife's age	8	Long-term, Short-term	{No-use}, {Long-term, Short-term}	{No-use}
Car.	Safety	5	acc, good, vgood	{unacc}, {acc, good, vgood}	{unacc}
Che.	wknck, wkovl, wkpos, wtoeg	32		{won}, {nowin}	{won}
Nur.	Health	7	recommend, very-recom, priority, spec-prior	{not-recom}, {recommend, priority, very-recom, spec-prior}	{not-recom}
Def.	Amount paid in June, amount paid in May, amount paid in April	20		{payment}, {nonpayment}	{payment}
Ban.	Age, balance, day, duration, campaign, pdays, previous	9		{non-subscription}, {subscription}	{non-subscription}
Con.	e_5, e_6 , $f_1, f_2, f_3, f_4, f_5, f_6, g_1, g_2, g_3, g_4, g_5, g_6$	28	{loss, draw}	{win}, {loss, draw}	{win}

Table 8
Data distribution information after pretreatment.

Data	Samples	Condition attributes	1st class (C)	2nd class
Tic.	958	6	626 (65.34%)	332 (34.66%)
Cmc.	1473	8	629 (42.70%)	844 (57.30%)
Car.	1728	5	1210 (70.02%)	518 (29.98%)
Che.	3196	32	1669 (52.22%)	1527 (47.78%)
Mus.	8124	16	4208 (51.80%)	3916 (48.20%)
Nur.	12960	7	4320 (33.33%)	8640 (66.67%)
Def.	30000	20	23364 (77.88%)	6636 (22.12%)
Ban.	45211	9	39992 (88.30%)	5289 (11.70%)
Con.	67557	28	44473 (65.83%)	23084 (34.17%)

Table 9
The granulation selection of data sets.

Data	A_1	A_2	A_3	A_4
Tic.	$\{a_1, a_3 - a_5\}$	$\{a_1, a_2, a_4, a_6\}$	$\{a_1, a_3, a_6\}$	$\{a_2 - a_6\}$
Cmc.	$\{a_1 - a_5\}$	$\{a_2 - a_6\}$	$\{a_3 - a_7\}$	$\{a_4 - a_8\}$
Car.	$\{a_1 - a_4\}$	$\{a_2 - a_5\}$	$\{a_1, a_3 - a_5\}$	$\{a_1, a_2, a_4, a_5\}$
Che.	$\{a_1 - a_8\}$	$\{a_9 - a_{16}\}$	$\{a_{17} - a_{24}\}$	$\{a_{25} - a_{32}\}$
Mus.	$\{a_1 - a_4\}$	$\{a_5 - a_8\}$	$\{a_9 - a_{12}\}$	$\{a_{13} - a_{16}\}$
Nur.	$\{a_1 - a_4\}$	$\{a_2 - a_5\}$	$\{a_3 - a_6\}$	$\{a_4 - a_7\}$
Def.	$\{a_1 - a_5\}$	$\{a_6 - a_{10}\}$	$\{a_{11} - a_{15}\}$	$\{a_{16} - a_{20}\}$
Ban.	$\{a_1 - a_3, a_8, a_9\}$	$\{a_2 - a_4, a_7, a_8\}$	$\{a_1, a_2, a_5 - a_7\}$	$\{a_3, a_4, a_6 - a_8\}$
Con.	$\{a_1 - a_7\}$	$\{a_8 - a_{14}\}$	$\{a_{15} - a_{21}\}$	$\{a_{22} - a_{28}\}$

of each granulation is obtained by normalizing in each data set. Detailed results are shown in Table 12. The granulation weights of the AAW, AQW, GEW methods are based on upper and lower approximations as described in section 2.4.

Secondly, under the interval-valued loss function, we propose the WGM-IVDTRSC, WGM-IVDTRSG and WGM-IVDTRSO models based on the different decision risk parameter determination methods. Next, we compare the classification error rates of these three models with those of other non-weighted multi-granulation decision models and other weighted models. In order to enhance the reliability of data results, we conducted five experiments on each data set by randomly selecting 20%, 40%, 60%, 80% and 100% of the data. In order to explain the advantage of the proposed granulation weighted method more objectively under the same decision risk parameters, we studied the classification error rate of each model under different information levels ($\varphi = 0.6, 0.8, 1.0$). At the same time, in order to understand the fluctuation of error rate with the increase of data size, the average classification error rate of each model is calculated in different data sets. The classification error rates of the WGM-IVDTRSC,

Table 10
The loss functions of different data sets.

Data	$\widetilde{\lambda}_{PP}$	$\widetilde{\lambda}_{PN}$	$\widetilde{\lambda}_{BP}$	$\widetilde{\lambda}_{BN}$	$\widetilde{\lambda}_{NP}$	$\widetilde{\lambda}_{NN}$
Tic.	[0.0510, 0.4701]	[0.6470, 0.9183]	[0.0837, 0.5880]	[0.3735, 0.8984]	[0.8259, 0.9830]	[0.0406, 0.5555]
Cmc.	[0.3074, 0.5015]	[0.4005, 0.5241]	[0.3944, 0.6047]	[0.2740, 0.3706]	[0.7369, 0.9248]	[0.2432, 0.3260]
Car.	[0.1896, 0.2533]	[0.8033, 0.9244]	[0.2176, 0.4168]	[0.6143, 0.6992]	[0.4391, 0.6352]	[0.2744, 0.6073]
Che.	[0.0871, 0.8280]	[0.5915, 0.7023]	[0.2329, 0.8875]	[0.3025, 0.6514]	[0.6520, 0.8945]	[0.2728, 0.4266]
Mus.	[0.1917, 0.3928]	[0.7014, 0.8570]	[0.2249, 0.5087]	[0.5163, 0.7606]	[0.4878, 0.9985]	[0.2197, 0.4928]
Nur.	[0.0560, 0.2931]	[0.2559, 0.5399]	[0.1184, 0.4097]	[0.0594, 0.4589]	[0.6117, 0.9810]	[0.0427, 0.0730]
Def.	[0.2484, 0.3523]	[0.5065, 0.6440]	[0.5249, 0.5407]	[0.1398, 0.4325]	[0.7370, 0.7491]	[0.1114, 0.3799]
Ban.	[0.0198, 0.2901]	[0.7308, 0.9473]	[0.2652, 0.4621]	[0.0398, 0.7457]	[0.5648, 0.7322]	[0.0335, 0.5973]
Con.	[0.0225, 0.2625]	[0.5673, 0.7207]	[0.0240, 0.4704]	[0.1055, 0.7198]	[0.0741, 0.7282]	[0.0905, 0.4746]

Table 11
The decision risk parameters of three methods.

Data	α_C	β_C	α_G	β_G	α_O	β_O
Tic.	0.6608	0.3727	0.5299	0.4357	0.8358	0.3257
Cmc.	0.5955	0.1021	0.6100	0.1431	0.8504	0.1571
Car.	0.6838	0.4954	0.5692	0.4576	0.7973	0.3624
Che.	0.6235	0.3739	0.5078	0.4073	0.8397	0.2494
Mus.	0.6538	0.4285	0.5568	0.4391	0.7911	0.3608
Nur.	0.6079	0.2744	0.5967	0.3648	0.8861	0.2539
Def.	0.5543	0.1614	0.5685	0.2446	0.8568	0.1491
Ban.	0.6814	0.2136	0.5800	0.2318	0.7614	0.2832
Con.	0.6884	0.4580	0.6689	0.3921	0.7469	0.3598

Table 12
The granulation weights based on the classification accuracy.

Tests	Tic.				Cmc.				Car.			
	acc_{A_1}	acc_{A_2}	acc_{A_3}	acc_{A_4}	acc_{A_1}	acc_{A_2}	acc_{A_3}	acc_{A_4}	acc_{A_1}	acc_{A_2}	acc_{A_3}	acc_{A_4}
1	0.6930	0.6763	0.5311	0.7718	0.6253	0.5758	0.6116	0.60885	0.6805	0.8138	0.6874	0.6736
2	0.5560	0.4730	0.3983	0.6473	0.6270	0.6297	0.6243	0.6108	0.4897	0.6552	0.6000	0.4713
3	0.5685	0.4357	0.4772	0.6017	0.6297	0.6027	0.5973	0.5324	0.4920	0.6989	0.6460	0.5402
4	0.6681	0.4730	0.6556	0.5519	0.6667	0.6198	0.6088	0.5840	0.6989	0.7770	0.6989	0.6989
ϖ	0.2708	0.2242	0.2247	0.2803	0.2613	0.2489	0.2503	0.2395	0.2287	0.2853	0.2550	0.2310
Tests	Che.				Mus.				Nur.			
	acc_{A_1}	acc_{A_2}	acc_{A_3}	acc_{A_4}	acc_{A_1}	acc_{A_2}	acc_{A_3}	acc_{A_4}	acc_{A_1}	acc_{A_2}	acc_{A_3}	acc_{A_4}
1	0.4938	0.5723	0.7107	0.5262	0.5992	0.5810	0.7647	0.6219	0.5444	0.6222	0.6667	0.6667
2	0.3591	0.4140	0.2681	0.3117	0.7922	0.9365	0.8949	0.9848	0.4180	0.5722	0.6667	0.6667
3	0.6160	0.5237	0.7045	0.4501	0.6755	0.9104	0.8897	0.9729	0.5667	0.6556	0.6667	0.6667
4	0.4434	0.6446	0.2320	0.3342	0.5022	0.8213	0.7775	0.8552	0.5944	0.6111	0.6667	0.6667
ϖ	0.2515	0.2833	0.2518	0.2133	0.2051	0.2594	0.2612	0.2742	0.2141	0.2481	0.2689	0.2689
Tests	Def.				Ban.				Con.			
	acc_{A_1}	acc_{A_2}	acc_{A_3}	acc_{A_4}	acc_{A_1}	acc_{A_2}	acc_{A_3}	acc_{A_4}	acc_{A_1}	acc_{A_2}	acc_{A_3}	acc_{A_4}
1	0.7788	0.8061	0.7821	0.7788	0.5266	0.2709	0.8598	0.2201	0.5995	0.6665	0.6337	0.6617
2	0.7788	0.8152	0.7849	0.7788	0.4834	0.3756	0.8825	0.3706	0.6472	0.5724	0.6819	0.6589
3	0.7787	0.8003	0.7784	0.7788	0.3957	0.3797	0.8829	0.3730	0.6851	0.6539	0.6514	0.6575
4	0.7788	0.7999	0.7812	0.7787	0.8831	0.8525	0.8830	0.8525	0.6851	0.6539	0.6514	0.6575
ϖ	0.2825	0.2826	0.2826	0.1523	0.2411	0.1979	0.3696	0.1913	0.2505	0.2401	0.2551	0.2543

WGM-IVDTRSG, WGM-IVDTRSO, MM, OM, PM, AAW, AQW and GEW models in case $\varphi = 0.6$ are given in detail. The remaining two cases mainly give the average classification error rates of the above models.

When $\varphi = 0.6$, the classification error rates of the WGM-IVDTRSC, MM, OM, PM, AAW, AQW and GEW models are shown in Table 13, where the underlined symbol represents the lowest classification error rate.

First, the relationships between the WGM-IVDTRSC model and other non-weighted models are analyzed. From Table 13, we find that the WGM-IVDTRSC model performs better than PM and OM on all data sets. In particular, on Cmc., Che., Mus. and Con. data sets, the WGM-IVDTRSC model performs far better than them. Compared with the MM model, on Cmc., Mus., and Con. data sets, the WGM-IVDTRSC model performs far better than MM. And on Car., Def. and Ban. data sets, the WGM-IVDTRSC model performs better than MM. On Che. and Nur. Data sets, the WGM-IVDTRSC model performs better than MM in most cases. Only on Tic data set, there are two experiments (20% and 80% of the data) the WGM-IVDTRSC model performs worse than MM; there is one experiment (60% of the data), the performances of the WGM-IVDTRSC model and MM are the same; and there are two experiments (40% and 100% of the data), the WGM-IVDTRSC model performs better than MM. To sum up, the WGM-IVDTRSC model based on the classification accuracy weighted method performs better than the non-weighted models MM, OM, PM. Therefore, different granulations are not equally important, and granulation weighted is necessary. Meanwhile, the relationships between the WGM-IVDTRSC model and other weighted models are analyzed. On seven data sets namely Tic., Car., Che., Mus., Nur., Ban. and Con., the WGM-IVDTRSC model performs better than AAW, AQW and GEW. On Cmc. data set, the WGM-IVDTRSC model performs better than AAW and AQW in most cases except the result of 60%, and the WGM-IVDTRSC model performs better than GEW. And on Def. data set, the performances of the WGM-IVDTRSC model, AAW and AQW are the same, and each of them is better than the performance of the GEW model. Therefore,

Table 13
The classification error rates under condition parameters (α_C, β_C) and $\varphi = 0.6$.

Tic.	WGM-IVDTRSC	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0157	0.0052	0.0419	0.0838	0.0209	0.0209	0.0419
40%	0.0236	0.0262	0.0785	0.0995	0.0471	0.0471	0.0759
60%	0.0279	0.0279	0.0836	0.0993	0.0453	0.0453	0.0627
80%	0.0405	0.0392	0.0941	0.0824	0.0471	0.0471	0.0444
100%	0.0282	0.0418	0.0981	0.0919	0.0407	0.0407	0.0564
Cmc.	WGM-IVDTRSC	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0068	0.0137	0.0410	0.0068	0.0079	0.0079	0.0098
40%	0.0034	0.0204	0.0323	0.0153	0.0058	0.0058	0.0170
60%	0.0091	0.0249	0.0294	0.0102	0.0088	0.0088	0.0181
80%	0.0051	0.0272	0.0535	0.0068	0.0065	0.0065	0.0263
100%	0.0054	0.0265	0.0563	0.0136	0.0054	0.0054	0.0238
Car.	WGM-IVDTRSC	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0087	0.0290	0.0638	0.1246	0.0098	0.0105	0.0209
40%	0.0362	0.0825	0.1085	0.1708	0.0662	0.0679	0.1324
60%	0.0676	0.0956	0.1245	0.1863	0.0782	0.0803	0.0876
80%	0.0709	0.1100	0.1375	0.1968	0.0963	0.1002	0.1028
100%	0.0799	0.1238	0.1319	0.1869	0.0944	0.1023	0.1236
Che.	WGM-IVDTRSC	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0031	0.0000	0.0657	0.1581	0.0082	0.0082	0.0188
40%	0.0039	0.0039	0.0681	0.1635	0.0072	0.0072	0.0141
60%	0.0016	0.0057	0.0610	0.1659	0.0088	0.0088	0.0136
80%	0.0016	0.0067	0.0626	0.1651	0.0073	0.0073	0.0133
100%	0.0009	0.0056	0.0638	0.1683	0.0069	0.0069	0.0100
Mus.	WGM-IVDTRSC	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0099	0.0074	0.0240	0.1435	0.0402	0.0402	0.1410
40%	0.0065	0.0105	0.0246	0.1391	0.0298	0.0298	0.1468
60%	0.0084	0.0113	0.0228	0.1399	0.0376	0.0376	0.1434
80%	0.0062	0.0109	0.0249	0.1647	0.0153	0.0153	0.1688
100%	0.0092	0.0137	0.0246	0.1620	0.0258	0.0258	0.1643
Nur.	WGM-IVDTRSC	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0123	0.0197	0.0498	0.1578	0.0246	0.0246	0.0623
40%	0.0058	0.0089	0.0226	0.1175	0.0134	0.0134	0.0356
60%	0.0010	0.0019	0.0015	0.0631	0.0477	0.0477	0.0921
80%	0.0008	0.0008	0.0008	0.0122	0.0098	0.0098	0.0765
100%	0.0000	0.0000	0.0000	0.0000	0.0147	0.0147	0.0327
Def.	WGM-IVDTRSC	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0007	0.0037	0.0030	0.0895	0.0007	0.0007	0.0025
40%	0.0012	0.0058	0.0048	0.0843	0.0012	0.0012	0.0047
60%	0.0011	0.0052	0.0034	0.0867	0.0011	0.0011	0.0047
80%	0.0011	0.0041	0.0055	0.0869	0.0011	0.0011	0.0040
100%	0.0011	0.0049	0.0027	0.0886	0.0011	0.0011	0.0038
Ban.	WGM-IVDTRSC	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0466	0.0709	0.0857	0.0491	0.0742	0.0742	0.0847
40%	0.0490	0.0788	0.0866	0.0553	0.0781	0.0781	0.0872
60%	0.0503	0.0805	0.0880	0.0569	0.0791	0.0791	0.0878
80%	0.0520	0.0819	0.0888	0.0596	0.0816	0.0816	0.0883
100%	0.0522	0.0828	0.0889	0.0596	0.0819	0.0819	0.0886
Con.	WGM-IVDTRSC	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0020	0.0210	0.0207	0.0774	0.0323	0.0323	0.0038
40%	0.0049	0.0233	0.0622	0.0787	0.0347	0.0347	0.0093
60%	0.0044	0.0233	0.0617	0.0806	0.0330	0.0330	0.0085
80%	0.0038	0.0233	0.0598	0.0834	0.0307	0.0307	0.0080
100%	0.0017	0.0237	0.0171	0.0844	0.0352	0.0352	0.0030

the WGM-IVDTRSC model based on the classification accuracy weighted method is feasible and effective.

In addition, the average classification error rates of the WGM-IVDTRSC, MM, OM, PM, AAW, AQW and GEW models under $\varphi =$

Table 14
The average classification error rates under parameters (α_C, β_C) and different φ.

Data	$\varphi = 0.6$						
	WGM-IVDTRSC	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
Tic.	0.0281 ± 0.0124	0.0234 ± 0.0182	0.0700 ± 0.0281	0.0909 ± 0.0085	0.0340 ± 0.0131	0.0340 ± 0.0131	0.0588 ± 0.0169
Cmc.	0.0062 ± 0.0028	0.0204 ± 0.0067	0.0428 ± 0.0134	0.0110 ± 0.0042	0.0071 ± 0.0017	0.0071 ± 0.0017	0.0180 ± 0.0082
Car.	0.0443 ± 0.0356	0.0764 ± 0.0474	0.1006 ± 0.0368	0.1607 ± 0.0361	0.0530 ± 0.0432	0.0564 ± 0.0459	0.0766 ± 0.0557
Che.	0.0024 ± 0.0015	0.0033 ± 0.0033	0.0645 ± 0.0035	0.1632 ± 0.0051	0.0078 ± 0.0009	0.0078 ± 0.0009	0.0144 ± 0.0044
Mus.	0.0080 ± 0.0018	0.0105 ± 0.0031	0.0238 ± 0.0010	0.1519 ± 0.0128	0.0277 ± 0.0124	0.0277 ± 0.0124	0.1549 ± 0.0139
Nur.	0.0061 ± 0.0061	0.0098 ± 0.0098	0.0248 ± 0.0248	0.0789 ± 0.0789	0.0287 ± 0.0189	0.0287 ± 0.0189	0.0624 ± 0.0297
Def.	0.0009 ± 0.0002	0.0047 ± 0.0010	0.0040 ± 0.0013	0.0868 ± 0.0025	0.0009 ± 0.0002	0.0009 ± 0.0002	0.0036 ± 0.0011
Ban.	0.0494 ± 0.0028	0.0768 ± 0.0059	0.0873 ± 0.0016	0.0543 ± 0.0052	0.0780 ± 0.0038	0.0780 ± 0.0038	0.0866 ± 0.0019
Con.	0.0032 ± 0.0015	0.0223 ± 0.0013	0.0396 ± 0.0225	0.0809 ± 0.0035	0.0329 ± 0.0022	0.0329 ± 0.0022	0.0061 ± 0.0031
Data	$\varphi = 0.8$						
	WGM-IVDTRSC	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
Tic.	0.0036 ± 0.0015	0.0234 ± 0.0182	0.0700 ± 0.0281	0.0909 ± 0.0085	0.0104 ± 0.0052	0.0104 ± 0.0052	0.0156 ± 0.0051
Cmc.	0.0017 ± 0.0017	0.0204 ± 0.0067	0.0428 ± 0.0134	0.0110 ± 0.0042	0.0034 ± 0.0019	0.0034 ± 0.0019	0.0068 ± 0.0034
Car.	0.0184 ± 0.0155	0.0764 ± 0.0474	0.1006 ± 0.0368	0.1607 ± 0.0361	0.0347 ± 0.0260	0.0184 ± 0.0155	0.0224 ± 0.0195
Che.	0.0000 ± 0.0000	0.0033 ± 0.0033	0.0645 ± 0.0035	0.1632 ± 0.0051	0.0015 ± 0.0006	0.0015 ± 0.0006	0.0048 ± 0.0023
Mus.	0.0000 ± 0.0000	0.0105 ± 0.0031	0.0238 ± 0.0010	0.1519 ± 0.0128	0.0035 ± 0.0031	0.0035 ± 0.0031	0.1549 ± 0.0139
Nur.	0.0015 ± 0.0015	0.0098 ± 0.0098	0.0248 ± 0.0248	0.0789 ± 0.0789	0.0051 ± 0.0051	0.0051 ± 0.0051	0.0134 ± 0.0061
Def.	0.0001 ± 0.0001	0.0047 ± 0.0010	0.004 ± 0.0013	0.0868 ± 0.0025	0.0001 ± 0.0001	0.0001 ± 0.0001	0.0003 ± 0.0001
Ban.	0.0331 ± 0.0032	0.0768 ± 0.0059	0.0873 ± 0.0016	0.0543 ± 0.0052	0.0540 ± 0.0051	0.0540 ± 0.0051	0.0743 ± 0.0041
Con.	0.0000 ± 0.0000	0.0223 ± 0.0013	0.0396 ± 0.0225	0.0809 ± 0.0035	0.0010 ± 0.0001	0.0010 ± 0.0001	0.0001 ± 0.0001
Data	$\varphi = 1.0$						
	WGM-IVDTRSC	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
Tic.	0.0036 ± 0.0015	0.0234 ± 0.0182	0.0700 ± 0.0281	0.0909 ± 0.0085	0.0036 ± 0.0015	0.0036 ± 0.0015	0.0036 ± 0.0015
Cmc.	0.0000 ± 0.0000	0.0204 ± 0.0067	0.0428 ± 0.0134	0.0110 ± 0.0042	0.0017 ± 0.0017	0.0017 ± 0.0017	0.0017 ± 0.0017
Car.	0.0184 ± 0.0155	0.0764 ± 0.0474	0.1006 ± 0.0368	0.1607 ± 0.0361	0.0184 ± 0.0155	0.0184 ± 0.0155	0.0184 ± 0.0155
Che.	0.0000 ± 0.0000	0.0033 ± 0.0033	0.0645 ± 0.0035	0.1632 ± 0.0051	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
Mus.	0.0000 ± 0.0000	0.0105 ± 0.0031	0.0238 ± 0.0010	0.1519 ± 0.0128	0.0000 ± 0.0000	0.0000 ± 0.0000	0.1549 ± 0.0139
Nur.	0.0015 ± 0.0015	0.0098 ± 0.0098	0.0248 ± 0.0248	0.0789 ± 0.0789	0.0015 ± 0.0015	0.0015 ± 0.0015	0.0015 ± 0.0015
Def.	0.0001 ± 0.0001	0.0047 ± 0.0010	0.0040 ± 0.0013	0.0868 ± 0.0025	0.0001 ± 0.0001	0.0001 ± 0.0001	0.0001 ± 0.0001
Ban.	0.0000 ± 0.0000	0.0768 ± 0.0059	0.0873 ± 0.0016	0.0543 ± 0.0052	0.0540 ± 0.0051	0.0000 ± 0.0000	0.0540 ± 0.0051
Con.	0.0000 ± 0.0000	0.0223 ± 0.0013	0.0396 ± 0.0225	0.0809 ± 0.0035	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000

0.6, 0.8, 1.0 are shown in Table 14, where the underlined symbol represents the lowest average classification error rate. From Table 14, when $\varphi = 0.6$, it is easy to find that the WGM-IVDTRSC model performs better than other models. Only on the Tic. data set, the performance of the WGM-IVDTRSC model is a little bit worse than that of the MM model. When $\varphi = 0.8$ and $\varphi = 1.0$, we find that the performance of the WGM-IVDTRSC model is the best on all the data sets. At the same time, we note that the average classification error rates of the four weighted models namely WGM-IVDTRSC, AAW, AQW and GEW models decrease with the increase of information levels. In particular, when the information level is equal to 1.0, the performances of the AAW, AQW and GEW models are greatly improved. Moreover, we find that the performance of the WGM-IVDTRSC model based on the classification accuracy weighted method is least affected by the changes of information levels, and it performs exceptionally well.

In order to be more intuitive, the relationships between the WGM-IVDTRSC method and other methods under different information levels are shown in Figs. 5–6. From Fig. 5, it is obvious that the WGM-IVDTRSC model performs very well when compared it with MM, OM, PM models. As can be seen from Fig. 6, the four weighted methods perform essentially the same when the information level is very high namely $\varphi = 1.0$. However, with the reduction of requirements, the performances of different weighted models are gradually reflected. With the decrease of information levels, the classification error rate is increasing. At the low information level, the WGM-IVDTRSC model performs well when compared it with the AAW, AQW and GEW weighted models.

Similarly, when $\varphi = 0.6$, the classification error rates of the WGM-IVDTRSG, MM, OM, PM, AAW, AQW and GEW models are shown in Table 15. And the average classification error rates of the WGM-IVDTRSG, MM, OM, PM, AAW, AQW and GEW models under different information levels are shown in Table 16. According to Tables 15–16, we find that the WGM-IVDTRSG model based on classification accuracy weighted method performs well when compared it with non-weighted models and weighted models under different information levels. For more intuitive information, see Figs. 7–8.

In the same way, the classification error rates of the WGM-IVDTRSO, MM, OM, PM, AAW, AQW and GEW models under $\varphi = 0.6$ are shown in Table 17. And the average classification error rates of the WGM-IVDTRSO, MM, OM, PM, AAW, AQW and GEW models under different information levels are shown in Table 18. According to the information in Tables 17–18, we know that the WGM-IVDTRSO performs well when compared it with other models under different information levels. Visual representation is shown in Figs. 9–10.

By comparing the proposed three WGM-IVDTRS models with the three non-weighted models MM, OM, PM, it is obvious that the weighted generalized multi-granulation interval-valued decision-theoretic rough set model based on the classification accuracy of decision tree learning is an important and feasible decision model for decision fusion. Moreover, by comparing the three WGM-IVDTRS model with the three weighted methods AAW, AQW, GEW, the feasibility and effectiveness of the WGM-IVDTRS model based on the classification accuracy of decision tree learning are verified.

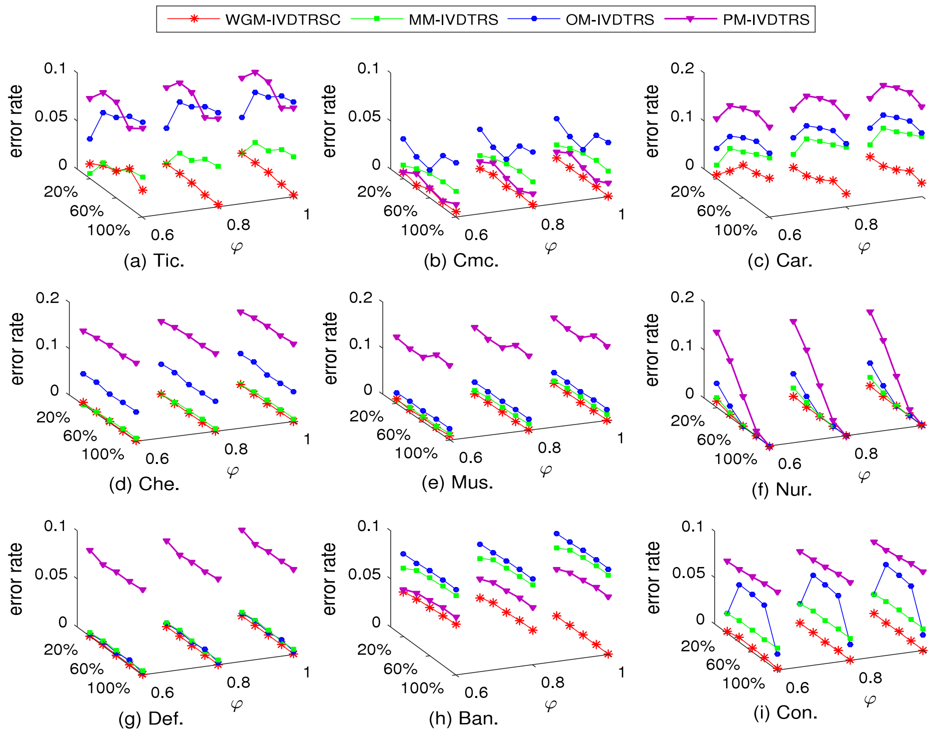


Fig. 5. Comparisons among WGM-IVDTRSC and three non-weighted models.

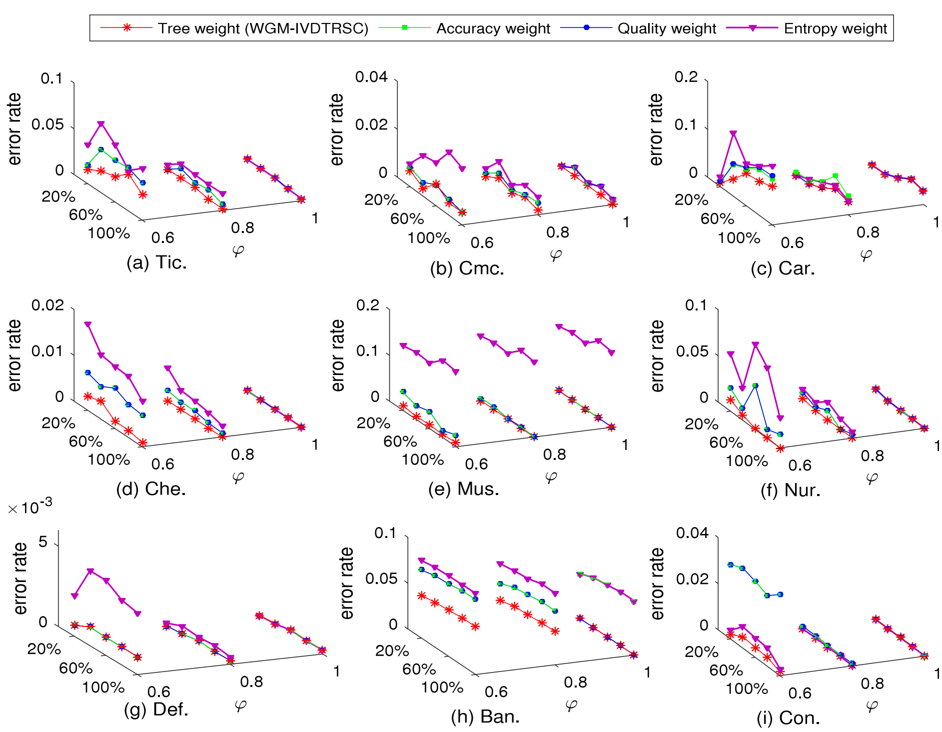


Fig. 6. Comparisons among WGM-IVDTRSC and three weighted models.

5.3. Comparison of three WGM-IVDTRS models

In order to evaluate the performances of the WGM-IVDTRSC, WGM-IVDTRSG and WGM-IVDTRSO models and find the criteria for selecting a suitable decision method under different environments, the classification error rates of decision classes are compared in the three models using the above nine data sets. First, the average classification error rates of the three WGM-IVDTRS models under different information levels are shown in Table 19, where the underline and the wavy line represent the smallest and the second smallest average classification error rates, respectively.

From Table 19, when $\varphi = 0.6$, the WGM-IVDTRSO model performs better than the WGM-IVDTRSC and WGM-IVDTRSG

Table 15
The classification error rates under condition parameters (α_G, β_G) and $\varphi = 0.6$.

Tic.	WGM-IVDTRSG	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0262	0.0524	0.0419	0.1152	0.0457	0.0457	0.1100
40%	0.0524	0.0838	0.1021	0.1728	0.0822	0.0822	0.1545
60%	0.1010	0.1272	0.1167	0.1882	0.1573	0.1573	0.1882
80%	0.0954	0.1294	0.1294	0.1621	0.1348	0.1348	0.1843
100%	0.0908	0.1326	0.1430	0.1649	0.0908	0.0908	0.1952
Cmc.	WGM-IVDTRSG	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0034	0.0000	0.0137	0.0068	0.0057	0.0057	0.0068
40%	0.0051	0.0051	0.0221	0.0119	0.0071	0.0071	0.0085
60%	0.0091	0.0045	0.0136	0.0091	0.0103	0.0103	0.0057
80%	0.0025	0.0068	0.0221	0.0034	0.0048	0.0048	0.0102
100%	0.0048	0.0061	0.0353	0.0115	0.0083	0.0083	0.0095
Car.	WGM-IVDTRSG	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0145	0.0377	0.0841	0.1304	0.0145	0.0145	0.0290
40%	0.0478	0.0825	0.1375	0.1751	0.0478	0.0478	0.0709
60%	0.0772	0.1043	0.1448	0.1873	0.0772	0.0772	0.0792
80%	0.0839	0.1201	0.1570	0.1975	0.0839	0.0839	0.0810
100%	0.0845	0.1395	0.1505	0.1869	0.0845	0.0845	0.0856
Che.	WGM-IVDTRSG	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0579	0.0876	0.0657	0.2034	0.0626	0.0626	0.0360
40%	0.0837	0.1009	0.1894	0.2113	0.1643	0.1643	0.1424
60%	0.0613	0.1127	0.4069	0.1894	0.1466	0.1466	0.2739
80%	0.0767	0.1080	0.1811	0.2140	0.1424	0.1424	0.1213
100%	0.0610	0.1105	0.1827	0.2566	0.1314	0.1314	0.1227
Mus.	WGM-IVDTRSG	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0197	0.0339	0.0240	0.1435	0.0213	0.0213	0.1410
40%	0.0098	0.0345	0.0400	0.1391	0.0147	0.0147	0.1468
60%	0.0218	0.0312	0.0382	0.1399	0.0352	0.0352	0.1434
80%	0.0089	0.0323	0.0405	0.1647	0.0112	0.0112	0.1688
100%	0.0220	0.0324	0.0414	0.1620	0.0378	0.0378	0.1643
Nur.	WGM-IVDTRSG	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.1316	0.1694	0.2010	0.3094	0.2467	0.2467	0.3346
40%	0.1931	0.2340	0.2593	0.3223	0.3442	0.3442	0.5874
60%	0.2469	0.2786	0.2967	0.3299	0.5614	0.5614	0.6037
80%	0.3040	0.3237	0.3232	0.3333	0.4589	0.4589	0.5573
100%	0.3333	0.3333	0.3333	0.3333	0.6777	0.6777	0.8814
Def.	WGM-IVDTRSG	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.1214	0.1235	0.2209	0.2212	0.1214	0.1214	0.1387
40%	0.1167	0.1239	0.2211	0.2212	0.1167	0.1167	0.1356
60%	0.1173	0.1244	0.2211	0.2212	0.1173	0.1173	0.1347
80%	0.1244	0.1260	0.2211	0.2212	0.1244	0.1244	0.1360
100%	0.1255	0.1267	0.2212	0.2212	0.1255	0.1255	0.1369
Ban.	WGM-IVDTRSG	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0561	0.0897	0.0960	0.0614	0.0919	0.0919	0.0958
40%	0.0577	0.0931	0.0967	0.0667	0.0927	0.0927	0.0971
60%	0.0578	0.0961	0.0923	0.0680	0.0915	0.0915	0.0942
80%	0.0584	0.0982	0.0954	0.0710	0.0925	0.0925	0.0958
100%	0.0590	0.0989	0.0931	0.0705	0.0926	0.0926	0.0949
Con.	WGM-IVDTRSG	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0392	0.0793	0.0702	0.0502	0.0885	0.0650	0.0412
40%	0.0370	0.0778	0.0684	0.0592	0.0852	0.0614	0.0389
60%	0.0390	0.0779	0.0684	0.0735	0.0865	0.0647	0.0396
80%	0.0388	0.0778	0.0688	0.0592	0.0865	0.0643	0.0394
100%	0.0383	0.0779	0.0683	0.0556	0.0862	0.0637	0.0390

models on all the data sets, and the WGM-IVDTRSC model performs better than the WGM-IVDTRSG model except the Cmc. data set. With the increase of parameter φ, the performances of the WGM-IVDTRSC and WGM-IVDTRSG models have been

Table 16
The average classification error rates under parameters (α_G, β_G) and different φ.

Data	WGM-IVDTRSG	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
$\varphi = 0.6$							
Tic.	0.0636 ± 0.0374	0.0924 ± 0.0400	0.0924 ± 0.0505	0.1517 ± 0.0365	0.1015 ± 0.0558	0.1015 ± 0.0558	0.1526 ± 0.0426
Cmc.	0.0058 ± 0.0033	0.0034 ± 0.0034	0.0244 ± 0.0108	0.0076 ± 0.0042	0.0075 ± 0.0027	0.0075 ± 0.0027	0.0079 ± 0.0022
Car.	0.0495 ± 0.0350	0.0886 ± 0.0509	0.1205 ± 0.0364	0.1639 ± 0.0335	0.0495 ± 0.0350	0.0495 ± 0.0350	0.0573 ± 0.0283
Che.	0.0707 ± 0.0128	0.1001 ± 0.0125	0.2363 ± 0.1706	0.2229 ± 0.0335	0.1134 ± 0.0508	0.1134 ± 0.0508	0.1549 ± 0.1189
Mus.	0.0154 ± 0.0065	0.0328 ± 0.0016	0.0327 ± 0.0087	0.1519 ± 0.0128	0.0245 ± 0.0133	0.0245 ± 0.0133	0.1549 ± 0.0139
Nur.	0.2324 ± 0.1008	0.2513 ± 0.0819	0.2671 ± 0.0661	0.3213 ± 0.0119	0.4621 ± 0.2154	0.4621 ± 0.2154	0.6079 ± 0.2733
Def.	0.1211 ± 0.0044	0.1251 ± 0.0016	0.2210 ± 0.0001	0.2212 ± 0.0000	0.1211 ± 0.0044	0.1211 ± 0.0044	0.1367 ± 0.0020
Ban.	0.0575 ± 0.0014	0.0943 ± 0.0046	0.0945 ± 0.0022	0.0661 ± 0.0047	0.0921 ± 0.0006	0.0921 ± 0.0006	0.0956 ± 0.0014
Con.	0.0381 ± 0.0011	0.0785 ± 0.0007	0.0692 ± 0.0009	0.0618 ± 0.0116	0.0868 ± 0.0016	0.0631 ± 0.0017	0.0400 ± 0.0011
$\varphi = 0.8$							
Tic.	0.0218 ± 0.0061	0.0924 ± 0.0400	0.0924 ± 0.0505	0.1517 ± 0.0365	0.0269 ± 0.0060	0.0218 ± 0.0061	0.0618 ± 0.0409
Cmc.	0.0000 ± 0.0000	0.0034 ± 0.0034	0.0244 ± 0.0108	0.0076 ± 0.0042	0.0013 ± 0.0013	0.0013 ± 0.0013	0.0029 ± 0.0021
Car.	0.0217 ± 0.0130	0.0886 ± 0.0509	0.1205 ± 0.0364	0.1639 ± 0.0335	0.0217 ± 0.0130	0.0217 ± 0.0130	0.0239 ± 0.0152
Che.	0.0086 ± 0.0086	0.1001 ± 0.0125	0.2363 ± 0.1706	0.2229 ± 0.0335	0.0211 ± 0.0133	0.0211 ± 0.0133	0.0160 ± 0.0144
Mus.	0.0000 ± 0.0000	0.0328 ± 0.0016	0.0327 ± 0.0087	0.1519 ± 0.0128	0.0000 ± 0.0000	0.0000 ± 0.0000	0.1549 ± 0.0139
Nur.	0.1853 ± 0.1479	0.2513 ± 0.0819	0.2671 ± 0.0661	0.3213 ± 0.0119	0.1853 ± 0.1479	0.1853 ± 0.1479	0.3015 ± 0.2563
Def.	0.0494 ± 0.0032	0.1251 ± 0.0016	0.2210 ± 0.0001	0.2212 ± 0.0000	0.0494 ± 0.0032	0.0494 ± 0.0032	0.0757 ± 0.0021
Ban.	0.0405 ± 0.0030	0.0943 ± 0.0046	0.0945 ± 0.0022	0.0661 ± 0.0047	0.0655 ± 0.0043	0.0655 ± 0.0043	0.0887 ± 0.0011
Con.	0.0029 ± 0.0006	0.0785 ± 0.0007	0.0692 ± 0.0009	0.0618 ± 0.0116	0.0175 ± 0.0007	0.0175 ± 0.0007	0.0168 ± 0.0003
$\varphi = 1.0$							
Tic.	0.0218 ± 0.0061	0.0924 ± 0.0400	0.0924 ± 0.0505	0.1517 ± 0.0365	0.0218 ± 0.0061	0.0218 ± 0.0061	0.0218 ± 0.0061
Cmc.	0.0000 ± 0.0000	0.0034 ± 0.0034	0.0244 ± 0.0108	0.0076 ± 0.0042	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
Car.	0.0217 ± 0.0130	0.0886 ± 0.0509	0.1205 ± 0.0364	0.1639 ± 0.0335	0.0217 ± 0.0130	0.0217 ± 0.0130	0.0217 ± 0.0130
Che.	0.0000 ± 0.0000	0.1001 ± 0.0125	0.2363 ± 0.1706	0.2229 ± 0.0335	0.0086 ± 0.0086	0.0000 ± 0.0000	0.0086 ± 0.0086
Mus.	0.0000 ± 0.0000	0.0328 ± 0.0016	0.0327 ± 0.0087	0.1519 ± 0.0128	0.0000 ± 0.0000	0.0000 ± 0.0000	0.1549 ± 0.0139
Nur.	0.1853 ± 0.1479	0.2513 ± 0.0819	0.2671 ± 0.0661	0.3213 ± 0.0119	0.1853 ± 0.1479	0.1853 ± 0.1479	0.1853 ± 0.1479
Def.	0.0494 ± 0.0032	0.1251 ± 0.0016	0.2210 ± 0.0001	0.2212 ± 0.0000	0.0494 ± 0.0032	0.0494 ± 0.0032	0.0494 ± 0.0032
Ban.	0.0000 ± 0.0000	0.0943 ± 0.0046	0.0945 ± 0.0022	0.0661 ± 0.0047	0.0655 ± 0.0043	0.0655 ± 0.0043	0.0655 ± 0.0043
Con.	0.0029 ± 0.0006	0.0785 ± 0.0007	0.0692 ± 0.0009	0.0618 ± 0.0116	0.0029 ± 0.0006	0.0029 ± 0.0006	0.0029 ± 0.0006

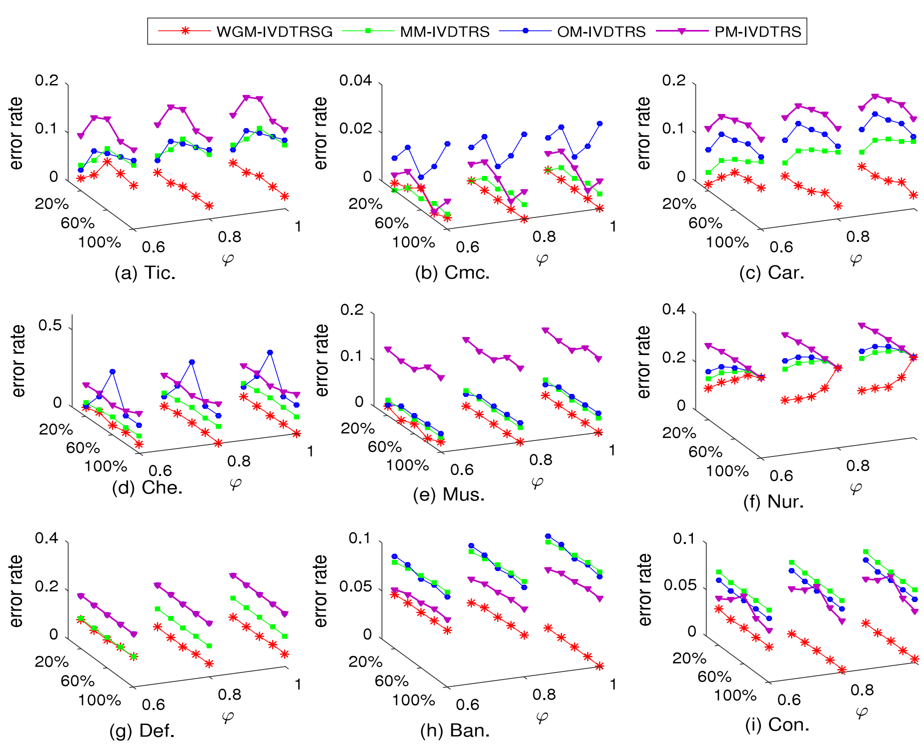


Fig. 7. Comparisons among WGM-IVDTRSG and three non-weighted models.

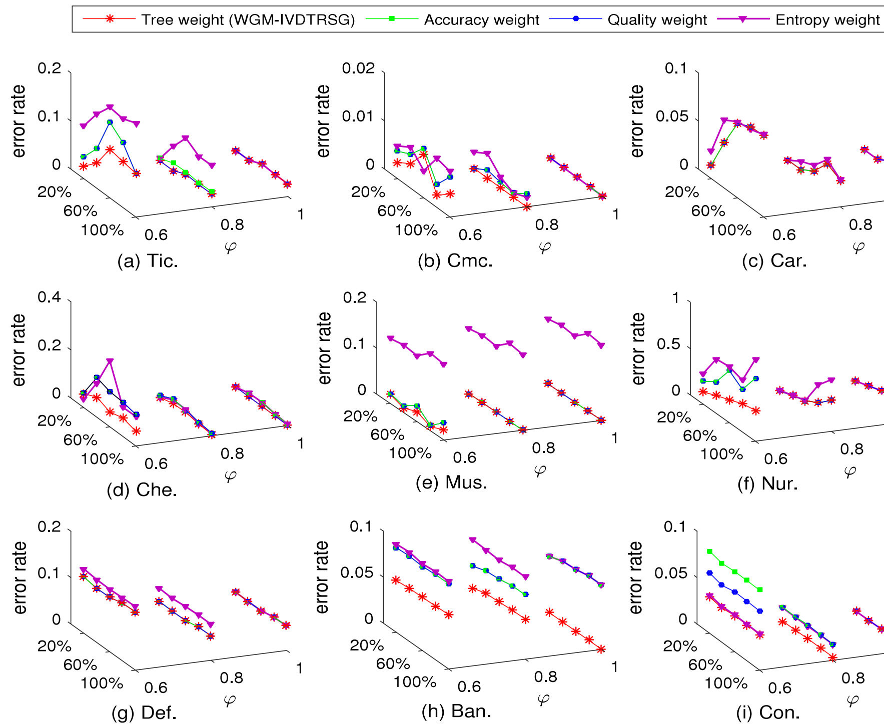


Fig. 8. Comparisons among WGM-IVDTRSG and three weighted models.

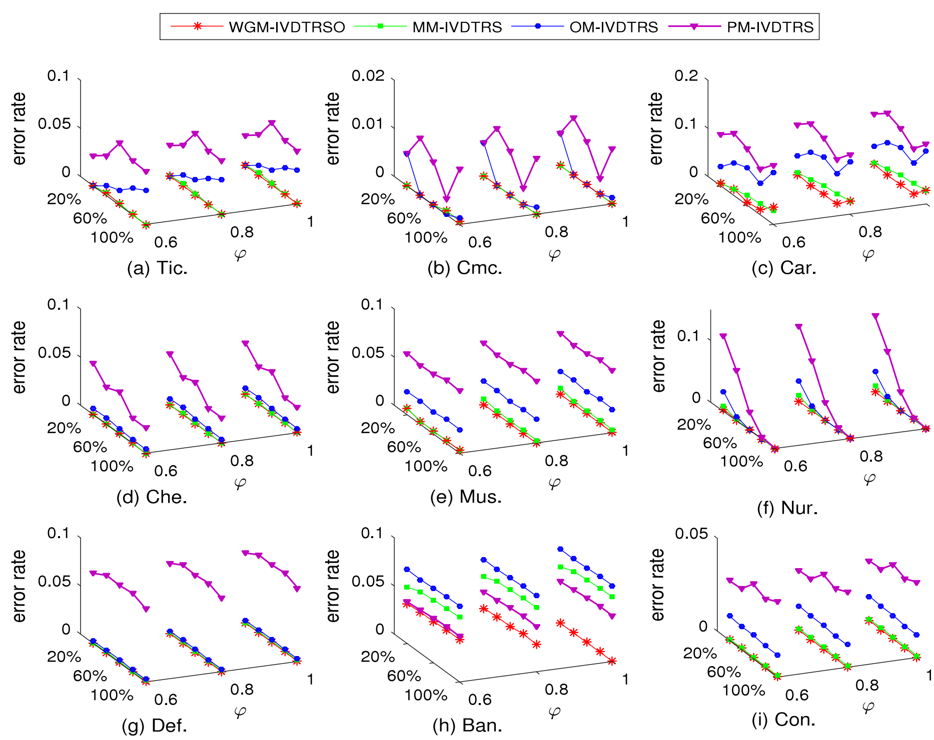


Fig. 9. Comparisons among WGM-IVDTRSO and three non-weighted models.

greatly improved. The WGM-IVDTRSO model performs better than the WGM-IVDTRSC and WGM-IVDTRSG models or is as good as them on all the data sets. What is more, the performance of the WGM-IVDTRSO model is least affected by the changes of information levels when compared it with the WGM-IVDTRSC and WGM-IVDTRSG models.

In order to be more intuitive, the experimental comparison results are shown in Figs. 11–12. It is important to note that some classification error rates are too small to be represented on the figures. Without affecting the logical size of the data, we add preprocessing to those too small data.

According to the experimental results, from the perspective of classification error rate, the WGM-IVDTRSO model has the best performance, the WGM-IVDTRSC model has better performance, and the performance of WGM-IVDTRSG model is worse than the other two models.

Table 17
The classification error rates under condition parameters (α_O, β_O) and $\varphi = 0.6$.

Tic.	WGM-IVDTRSO	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0000	0.0000	0.0000	0.0314	0.0142	0.0142	0.0105
40%	0.0026	0.0026	0.0105	0.0419	0.0135	0.0135	0.0137
60%	0.0017	0.0017	0.0157	0.0645	0.0189	0.0189	0.0000
80%	0.0000	0.0000	0.0275	0.0562	0.0147	0.0147	0.0098
100%	0.0000	0.0000	0.0355	0.0553	0.0279	0.0279	0.0000
Cmc.	WGM-IVDTRSO	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0000	0.0000	0.0068	0.0068	0.0134	0.0134	0.0114
40%	0.0000	0.0000	0.0000	0.0119	0.0117	0.0117	0.0085
60%	0.0000	0.0000	0.0000	0.0091	0.0211	0.0211	0.0000
80%	0.0008	0.0008	0.0000	0.0034	0.0039	0.0039	0.0019
100%	0.0007	0.0007	0.0014	0.0115	0.0050	0.0050	0.0026
Car.	WGM-IVDTRSO	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0058	0.0058	0.0406	0.1073	0.0173	0.0173	0.0058
40%	0.0130	0.0130	0.0680	0.1288	0.0389	0.0389	0.0130
60%	0.0068	0.0203	0.0782	0.1178	0.0129	0.0129	0.0068
80%	0.0101	0.0224	0.0644	0.0948	0.0243	0.0243	0.0101
100%	0.0370	0.0278	0.1088	0.1238	0.0567	0.0567	0.0370
Che.	WGM-IVDTRSO	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0000	0.0000	0.0063	0.0532	0.0000	0.0000	0.0031
40%	0.0008	0.0008	0.0070	0.0383	0.0156	0.0156	0.0078
60%	0.0005	0.0005	0.0052	0.0438	0.0313	0.0313	0.0016
80%	0.0004	0.0004	0.0043	0.0262	0.0078	0.0078	0.0016
100%	0.0000	0.0000	0.0041	0.0263	0.0063	0.0063	0.0013
Mus.	WGM-IVDTRSO	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0062	0.0062	0.0240	0.0640	0.0199	0.0199	0.0089
40%	0.0034	0.0034	0.0246	0.0616	0.0211	0.0211	0.0149
60%	0.0031	0.0031	0.0228	0.0624	0.0199	0.0199	0.0057
80%	0.0031	0.0031	0.0249	0.0659	0.0219	0.0219	0.0077
100%	0.0030	0.0030	0.0246	0.0645	0.0218	0.0218	0.0064
Nur.	WGM-IVDTRSO	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0039	0.0093	0.0324	0.1227	0.0189	0.0189	0.0089
40%	0.0004	0.0015	0.0071	0.0810	0.0057	0.0057	0.0027
60%	0.0000	0.0000	0.0000	0.0291	0.0135	0.0135	0.0000
80%	0.0000	0.0000	0.0000	0.0029	0.0097	0.0097	0.0017
100%	0.0000	0.0000	0.0000	0.0000	0.0048	0.0048	0.0000
Def.	WGM-IVDTRSO	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0005	0.0007	0.0025	0.0730	0.0495	0.0495	0.0013
40%	0.0008	0.0013	0.0030	0.0810	0.0645	0.0645	0.0033
60%	0.0009	0.0013	0.0034	0.0805	0.0638	0.0638	0.0034
80%	0.0007	0.0011	0.0025	0.0823	0.0647	0.0647	0.0028
100%	0.0008	0.0013	0.0027	0.0765	0.0628	0.0628	0.0027
Ban.	WGM-IVDTRSO	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0414	0.0586	0.0769	0.0432	0.0655	0.0655	0.0767
40%	0.0428	0.0641	0.0766	0.0448	0.0675	0.0675	0.0775
60%	0.0429	0.0653	0.0776	0.0465	0.0672	0.0672	0.0775
80%	0.0441	0.0664	0.0786	0.0477	0.0695	0.0695	0.0803
100%	0.0440	0.0670	0.0790	0.0473	0.0684	0.0684	0.0796
Con.	WGM-IVDTRSO	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
20%	0.0002	0.0007	0.0132	0.0322	0.0019	0.0019	0.0015
40%	0.0007	0.0010	0.0126	0.0328	0.0059	0.0059	0.0018
60%	0.0006	0.0009	0.0128	0.0402	0.0057	0.0057	0.0015
80%	0.0005	0.0009	0.0119	0.0374	0.0062	0.0062	0.0014
100%	0.0005	0.0009	0.0121	0.0407	0.0064	0.0064	0.0012

Table 18
The average classification error rates under parameters (α_O, β_O) and different φ.

Data	$\varphi = 0.6$						
	WGM-IVDTRSO	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
Tic.	0.0013 ± 0.0013	0.0013 ± 0.0013	0.0177 ± 0.0177	0.0479 ± 0.0165	0.0207 ± 0.0072	0.0207 ± 0.0072	0.0068 ± 0.0068
Cmc.	0.0004 ± 0.0004	0.0004 ± 0.0004	0.0034 ± 0.0034	0.0076 ± 0.0042	0.0125 ± 0.0086	0.0125 ± 0.0086	0.0057 ± 0.0057
Car.	0.0214 ± 0.0156	0.0168 ± 0.0110	0.0747 ± 0.0341	0.1118 ± 0.0170	0.0348 ± 0.0219	0.0348 ± 0.0219	0.0214 ± 0.0156
Che.	0.0004 ± 0.0004	0.0004 ± 0.0004	0.0055 ± 0.0014	0.0396 ± 0.0134	0.0156 ± 0.0156	0.0156 ± 0.0156	0.0045 ± 0.0032
Mus.	0.0045 ± 0.0015	0.0045 ± 0.0015	0.0238 ± 0.0010	0.0637 ± 0.0021	0.0208 ± 0.0009	0.0208 ± 0.0009	0.0103 ± 0.0046
Nur.	0.0019 ± 0.0019	0.0046 ± 0.0046	0.0162 ± 0.0162	0.0613 ± 0.0613	0.0118 ± 0.0070	0.0118 ± 0.0070	0.0044 ± 0.0044
Def.	0.0006 ± 0.0001	0.0009 ± 0.0002	0.0029 ± 0.0004	0.0776 ± 0.0046	0.0570 ± 0.0075	0.0570 ± 0.0075	0.0023 ± 0.0010
Ban.	0.0427 ± 0.0013	0.0628 ± 0.0042	0.0777 ± 0.0011	0.0454 ± 0.0022	0.0675 ± 0.0020	0.0675 ± 0.002	0.0784 ± 0.0017
Con.	0.0004 ± 0.0002	0.0008 ± 0.0001	0.0125 ± 0.0006	0.0364 ± 0.0042	0.0041 ± 0.0022	0.0041 ± 0.0022	0.0015 ± 0.0003
Data	$\varphi = 0.8$						
	WGM-IVDTRSO	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
Tic.	0.0000 ± 0.0000	0.0013 ± 0.0013	0.0177 ± 0.0177	0.0479 ± 0.0165	0.0013 ± 0.0013	0.0177 ± 0.0177	0.0479 ± 0.0165
Cmc.	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0034 ± 0.0034	0.0076 ± 0.0042	0.0066 ± 0.0066	0.0066 ± 0.0066	0.0000 ± 0.0000
Car.	0.0139 ± 0.0139	0.0168 ± 0.0110	0.0747 ± 0.0341	0.1118 ± 0.0170	0.0265 ± 0.0265	0.0265 ± 0.0265	0.0139 ± 0.0139
Che.	0.0000 ± 0.0000	0.0004 ± 0.0004	0.0055 ± 0.0014	0.0396 ± 0.0134	0.0004 ± 0.0004	0.0004 ± 0.0004	0.0004 ± 0.0004
Mus.	0.0000 ± 0.0000	0.0045 ± 0.0015	0.0238 ± 0.0010	0.0637 ± 0.0021	0.0024 ± 0.0010	0.0024 ± 0.0010	0.0057 ± 0.0006
Nur.	0.0000 ± 0.0000	0.0046 ± 0.0046	0.0162 ± 0.0162	0.0613 ± 0.0613	0.0011 ± 0.0011	0.0011 ± 0.0011	0.0020 ± 0.0020
Def.	0.0001 ± 0.0001	0.0009 ± 0.0002	0.0029 ± 0.0004	0.0776 ± 0.0046	0.0068 ± 0.0016	0.0083 ± 0.0021	0.0004 ± 0.0001
Ban.	0.0277 ± 0.0014	0.0628 ± 0.0042	0.0777 ± 0.0011	0.0454 ± 0.0022	0.0438 ± 0.0021	0.0438 ± 0.0021	0.0591 ± 0.0006
Con.	0.0000 ± 0.0000	0.0008 ± 0.0001	0.0125 ± 0.0006	0.0364 ± 0.0042	0.0002 ± 0.0001	0.0002 ± 0.0001	0.0000 ± 0.0000
Data	$\varphi = 1.0$						
	WGM-IVDTRSO	Non-weighted models			Other weighted models		
		MM	OM	PM	AAW	AQW	GEW
Tic.	0.0000 ± 0.0000	0.0013 ± 0.0013	0.0177 ± 0.0177	0.0479 ± 0.0165	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
Cmc.	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0034 ± 0.0034	0.0076 ± 0.0042	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
Car.	0.0139 ± 0.0139	0.0168 ± 0.0110	0.0747 ± 0.0341	0.1118 ± 0.0170	0.0139 ± 0.0139	0.0139 ± 0.0139	0.0139 ± 0.0139
Che.	0.0000 ± 0.0000	0.0004 ± 0.0004	0.0055 ± 0.0014	0.0396 ± 0.0134	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
Mus.	0.0000 ± 0.0000	0.0045 ± 0.0015	0.0238 ± 0.0010	0.0637 ± 0.0021	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
Nur.	0.0000 ± 0.0000	0.0046 ± 0.0046	0.0162 ± 0.0162	0.0613 ± 0.0613	0.0005 ± 0.0005	0.0005 ± 0.0005	0.0005 ± 0.0005
Def.	0.0001 ± 0.0001	0.0009 ± 0.0002	0.0029 ± 0.0004	0.0776 ± 0.0046	0.0001 ± 0.0001	0.0001 ± 0.0001	0.0001 ± 0.0001
Ban.	0.0000 ± 0.0000	0.0628 ± 0.0042	0.0777 ± 0.0011	0.0454 ± 0.0022	0.0000 ± 0.0000	0.0438 ± 0.0021	0.0438 ± 0.0021
Con.	0.0000 ± 0.0000	0.0008 ± 0.0001	0.0125 ± 0.0006	0.0364 ± 0.0042	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000

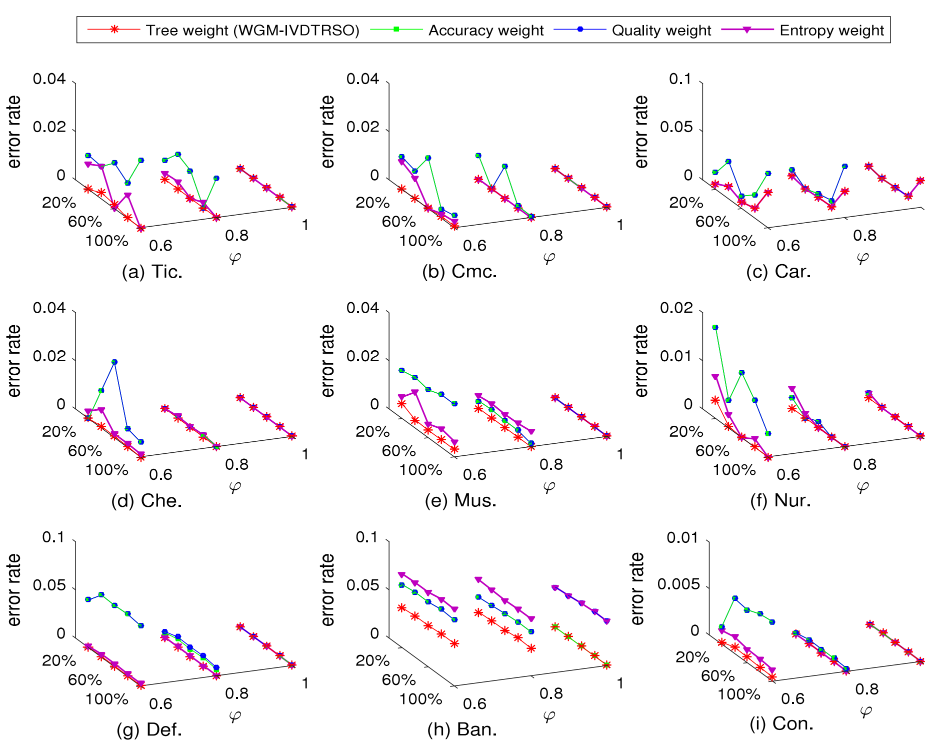


Fig. 10. Comparisons among WGM-IVDTRSO and three weighted models.

Table 19
The average error rates of three WGM-IVDTRS.

Data	$\varphi = 0.6$		
	WGM-IVDTRSC	WGM-IVDTRSG	WGM-IVDTRSO
Tic.	0.0281 ± 0.0124	0.0636 ± 0.0374	0.0013 ± 0.0013
Cmc.	0.0062 ± 0.0028	0.0058 ± 0.0033	0.0004 ± 0.0004
Car.	0.0443 ± 0.0356	0.0495 ± 0.0350	0.0214 ± 0.0156
Che.	0.0024 ± 0.0015	0.0707 ± 0.0128	0.0004 ± 0.0004
Mus.	0.0080 ± 0.0018	0.0154 ± 0.0065	0.0045 ± 0.0015
Nur.	0.0061 ± 0.0061	0.2324 ± 0.1008	0.0019 ± 0.0019
Def.	0.0009 ± 0.0002	0.1211 ± 0.0044	0.0006 ± 0.0001
Ban.	0.0494 ± 0.0028	0.0575 ± 0.0014	0.0427 ± 0.0013
Con.	0.0032 ± 0.0015	0.0381 ± 0.0011	0.0004 ± 0.0002
Data	$\varphi = 0.8$		
	WGM-IVDTRSC	WGM-IVDTRSG	WGM-IVDTRSO
Tic.	0.0036 ± 0.0015	0.0218 ± 0.0061	0.0000 ± 0.0000
Cmc.	0.0017 ± 0.0017	0.0000 ± 0.0000	0.0000 ± 0.0000
Car.	0.0184 ± 0.0155	0.0217 ± 0.0130	0.0139 ± 0.0139
Che.	0.0000 ± 0.0000	0.0086 ± 0.0086	0.0000 ± 0.0000
Mus.	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
Nur.	0.0015 ± 0.0015	0.1853 ± 0.1479	0.0000 ± 0.0000
Def.	0.0001 ± 0.0001	0.0494 ± 0.0032	0.0001 ± 0.0001
Ban.	0.0331 ± 0.0032	0.0405 ± 0.003	0.0277 ± 0.0014
Con.	0.0000 ± 0.0000	0.0029 ± 0.0006	0.0000 ± 0.0000
Data	$\varphi = 1.0$		
	WGM-IVDTRSC	WGM-IVDTRSG	WGM-IVDTRSO
Tic.	0.0036 ± 0.0015	0.0218 ± 0.0061	0.0000 ± 0.0000
Cmc.	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
Car.	0.0184 ± 0.0155	0.0217 ± 0.0130	0.0139 ± 0.0139
Che.	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
Mus.	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
Nur.	0.0015 ± 0.0015	0.1853 ± 0.1479	0.0000 ± 0.0000
Def.	0.0001 ± 0.0001	0.0494 ± 0.0032	0.0001 ± 0.0001
Ban.	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
Con.	0.0000 ± 0.0000	0.0029 ± 0.0006	0.0000 ± 0.0000

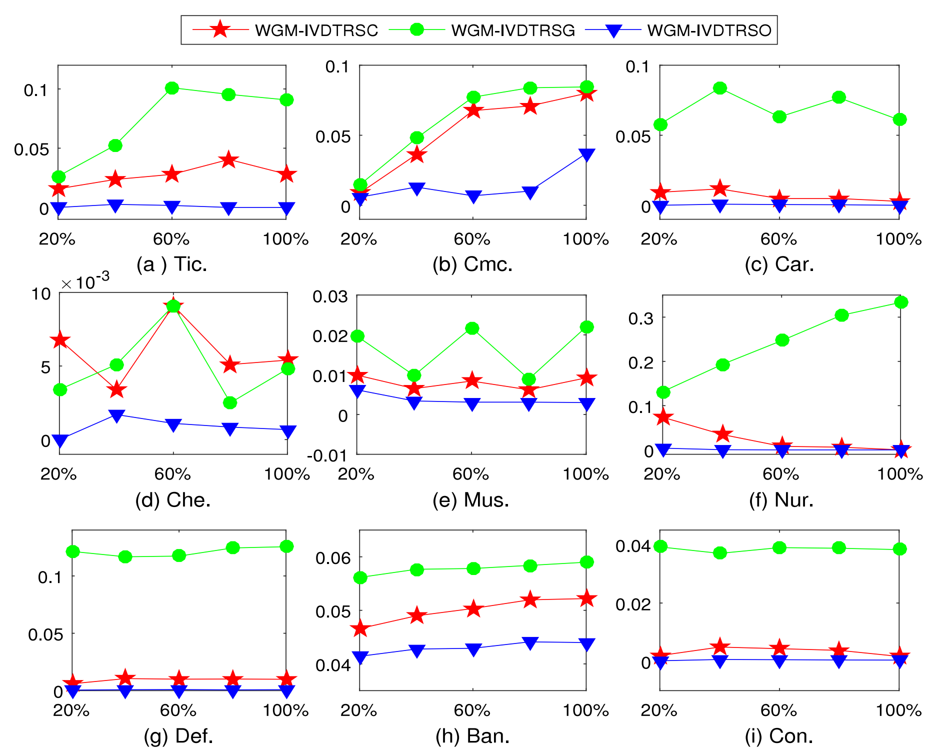


Fig. 11. Comparisons among WGM-IVDTRSC, WGM-IVDTRSG and WGM-IVDTRSO under $\varphi = 0.6$.

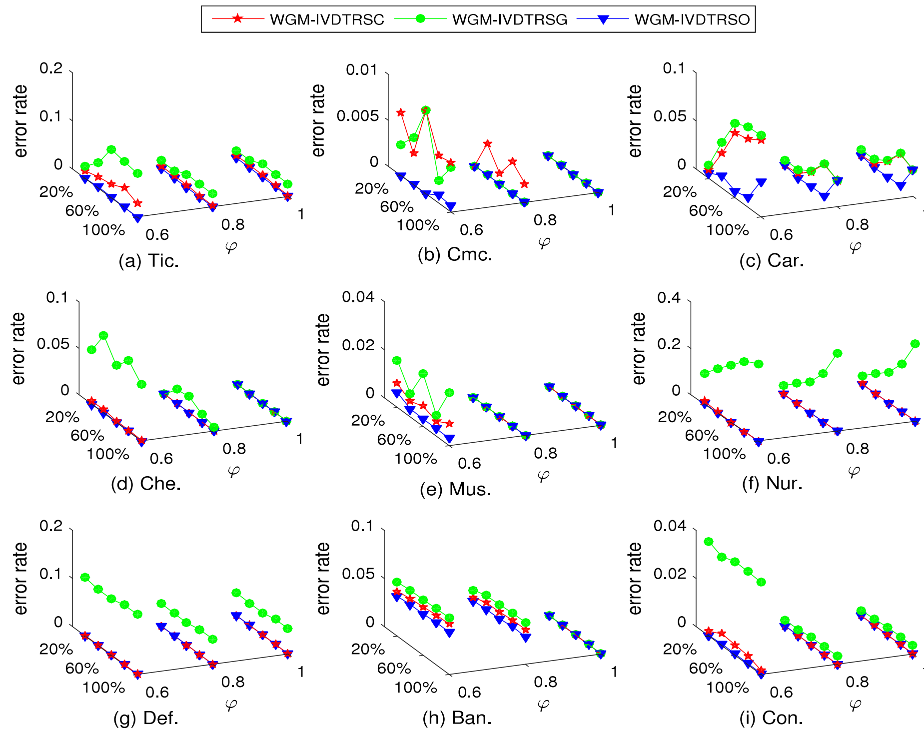


Fig. 12. Comparisons among three WGM-IVDTRS models under different information levels φ.

5.4. Experimental summary

The classification accuracy weighted method is a feasible and effective granulation weighted method from the point of view of classification error rate by comparing it with different weighted methods. And the weighted generalized multi-granulation interval-valued decision-theoretic rough set models based on the classification accuracy weighted method are important decision models. They provide methods for multi-granulation fusion and decision making. According to the comparison results of three WGM-IVDTRS models, we can choose different models in different circumstances. When decision makers can accept any value of an interval that can be used to represent the interval, the WGM-IVDTRSO model based on optimization method has the lowest classification error rate. This means that when people accept interval numbers as the extensibility and fault tolerance of information granulation, the WGM-IVDTRSO model is our best choice. Then for the moderate risk preference, the WGM-IVDTRSC model and WGM-IVDTRSG can meet our needs. And the WGM-IVDTRSC model is a relatively good choice. Of course, the WGM-IVDTRSG model can also be used as a decision analysis method in situations where certain error rate can be accepted.

The WGM-IVDTRS models in this paper mainly focus on the determination of decision risk parameters α, β and the granulation weighted method. The influence of information levels on model classification performance can be further studied. How to choose the information level, i.e. the number of granulations, is also an important research topic under certain model classification requirements. The influences of decision risk parameters, information levels and granulation weights on the performance of the model need further systematic study.

6. Conclusions

Weighted multi-granulation rough set theory as a multi-view data analysis method can effectively mine knowledge from data. Therefore, this paper introduces weighted generalized multi-granulation rough sets into the decision-theoretic rough set model to explore decision making problems of multi-source decision systems with different attribute sets. Considering the limitation of the actual conditions and the imprecision of expert evaluations, in this paper we express the loss function with intervals. The main work of this paper is to propose a new granulation weighted method based on the classification accuracy of decision tree learning from the point of view of machine learning, and construct the basic form of weighted generalized multi-granulation interval-valued decision-theoretic rough sets (WGM-IVDTRS). Meanwhile, three types of the WGM-IVDTRS model are proposed based on three different parameter determination methods. Finally, three WGM-IVDTRS models are proved to be an important, feasible and effective decision models by comparing them with other multi-granulation decision models. And the third type of the WGM-IVDTRS model (WGM-IVDTRSO) is found to perform best in terms of classification error rate by numerical experimental analysis when people accept the range scalability and fault tolerance of intervals. Therefore, it is meaningful to combine human cognitive ability with machine learning ability. In the future, based on multi-granulation theory and machine learning methods (fuzzy decision tree, recurrent neural network and deep neural network), we will study information fusion and decision-making of multi-source fuzzy data, multi-source dynamic data and multi-modal data. By combining the respective advantages of machine learning methods and multi-granulation theory, new models can be developed to accommodate more complex data environment of contemporary times, which can help people mine data and discover knowledge.

Acknowledgments

This work is supported by the Macau Science and Technology Development Fund (no. 081/2015/A3), the Natural Science Foundation of China (no. 61772002). We would like to thank the editor, associate editor and reviewers for their valuable suggestions.

References

[1] L. Zadeh, Fuzzy logic=computing with words, IEEE Trans. Fuzzy Syst. 4 (2) (1996) 103–111.
[2] Z. Pawlak, Rough sets, Int. J. Comput. Inf. Sci. 11 (5) (1982) 341–356.
[3] Y.H. Qian, J.Y. Liang, Y.Y. Yao, C.Y. Dang, MGRS: A multi-granulation rough set, Inform. Sci. 180 (6) (2010) 949–970.
[4] Y.Y. Yao, S.K.M. Wong, A decision theoretic framework for approximating concepts, Int. J. Man-Mach. Stud. 37 (6) (1992) 793–809.
[5] Y.Y. Yao, S.K.M. Wong, P. Lingras, A decision-theoretic rough set model, in: Z.W. Ras, M. Zemankova, M.L. Emrich (Eds.), in: Methodologies for Intelligent Systems, vol. 5, North-Holland, New York, 1990.
[6] Y.Y. Yao, Decision-theoretic rough set models, in: International Conference on Rough Sets and Knowledge Technology, 2007, pp. 1–12.
[7] Y.Y. Yao, The superiority of three-way decisions in probabilistic rough set models, Inform. Sci. 181 (6) (2011) 1080–1096.
[8] Y.Y. Yao, Y.H. She, Rough set models in multigranulation spaces, Inform. Sci. 327 (2016) 40–56.
[9] D. Liu, Y.Y. Yao, T.R. Li, Three-way investment decisions with decision-theoretic rough sets, Int. J. Comput. Intell. Syst. 4 (1) (2011) 66–74.
[10] D. Liu, T.R. Li, D.C. Liang, Three-way government decision analysis with decision-theoretic rough sets, Int. J. Uncertainty Fuzz. Knowl.-Based Syst. 20 (1) (2012) 119–132.
[11] Y. Zhang, J.T. Yao, Determining three-way decision regions with gini coefficients, in: International Conference on Rough Sets and Current Trends in Computing, Springer, 2014, pp. 160–171.
[12] B. Zhou, Y.Y. Yao, J.G. Luo, Cost-sensitive three-way email spam filtering, J. Intell. Inf. Syst. 42 (1) (2014) 19–45.
[13] Z.T. Gong, R.L. Chai, Covering multigranulation trapezoidal fuzzy decision-theoretic rough fuzzy set models and applications, J. Intell. Fuzzy Syst. 31 (3) (2016) 1621–1633.
[14] D. Liu, T.R. Li, D.C. Liang, Fuzzy decision-theoretic rough sets, Comput. Sci. 39 (12) (2012) 25–29.
[15] D. Liu, T.R. Li, H.X. Li, Interval-valued decision-theoretic rough sets, Comput. Sci. 39 (7) (2012) 178–214.
[16] D.C. Liang, D. Liu, Systematic studies on three-way decisions with interval-valued decision-theoretic rough sets, Inform. Sci. 276 (2014) 186–203.
[17] H.D. Zhao, Y.E. Bao, S.X. Guan, Geometry average sorting function about interval number and its application, J. Inner Mongolia Univ. Natl. (2011).
[18] Y.H. Qian, H. Zhang, Y.L. Sang, J.Y. Liang, Multi-granulation decision-theoretic rough sets, Internat. J. Approx. Reason. 55 (1) (2014) 225–237.
[19] T. Feng, J.S. Mi, Variable precision multi-granulation decision-theoretic fuzzy rough sets, Knowl.-Based Syst. 91 (2016) 93–101.
[20] J.H. Yu, B. Zhang, M.H. Chen, W.H. Xu, Double-quantitative decision-theoretic approach to multigranulation approximate space, Internat. J. Approx. Reason. 98 (2018) 236–258.
[21] B.Z. Sun, W.M. Ma, X. Xiao, Three-way group decision making based on multi-granulation fuzzy decision-theoretic rough set over two universes, Internat. J. Approx. Reason. 81 (2017) 87–102.
[22] W.T. Li, W.H. Xu, Multi-granulation decision-theoretic rough set in ordered information system, Fund. Inform. 139 (1) (2015) 67–89.
[23] C.H. Liu, W. Pedrycz, M.Z. Wang, Covering-based multigranulation decision-theoretic rough sets, J. Intell. Fuzzy Syst. 32 (1) (2017) 749–765.
[24] G.P. Lin, J.Y. Liang, Y.H. Qian, J.J. Li, A fuzzy multi-granulation decision-theoretic approach to multi-source fuzzy information systems, Knowl.-Based Syst. 91 (2016) 102–113.
[25] Z.W. Li, X.F. Liu, G.Q. Zhang, N.X. Xie, S.C. Wang, A multi-granulation decision-theoretic rough set method for distributed fc-decision information systems: An application in medical diagnosis, Appl. Soft Comput. 56 (2017) 233–244.
[26] J.M. Zhan, B.Z. Sun, J. Carlos, R. Alcantud, Covering based multigranulation $(\mathcal{I}, \mathcal{T})$−fuzzy rough set models and applications in multi-attribute group decision-making, Inform. Sci. 476 (2019) 290–318.
[27] H.B. Jiang, J.M. Zhan, D.G. Chen, Covering based variable precision $(\mathcal{I}, \mathcal{T})$−fuzzy rough sets with applications to multi-attribute decision-making, IEEE Trans. Fuzzy Syst. (2018).
[28] W.H. Xu, Y.T. Guo, Generalized multi-granulation double-quantitative decision-theoretic rough set, Knowl.-Based Syst. 105 (2016) 190–205.
[29] Y.T. Guo, E.C.C. Tsang, W.H. Xu, A weighted multi-granulation decision-theoretic approach to multi-source decision systems, in: International Conference on Machine Learning and Cybernetics, IEEE, 2017, pp. 202–210.
[30] X. Ji, P. Zhao, S. Yao, Rough set model and decision research in intuitionistic fuzzy information system based on weighted multi-granulation, Pattern Recog. Artif. Intell. 30 (11) (2017) 971–982.
[31] W.H. Xu, Ordered Information Systems and Rough Sets, Science Press, Beijing, 2013.
[32] D.G. Chen, S.Y. Zhao, L. Zhang, Y.P. Yang, X. Zhang, Sample pair selection for attribute reduction with rough set, IEEE Trans. Knowl. Data Eng. 24 (11) (2012) 2080–2093.
[33] E.C.C. Tsang, D.G. Chen, D.S. Yeung, X.Z. Wang, Attributes reduction using fuzzy rough sets, IEEE Trans. Fuzzy Syst. 16 (5) (2008) 1130–1141.
[34] Y.Y. Yang, D.G. Chen, H. Wang, X.Z. Wang, Incremental perspective for feature selection based on fuzzy rough sets, IEEE Trans. Fuzzy Syst. 26 (3) (2018) 1257–1273.
[35] D. Liu, T.R. Li, D. Ruan, W.L. Zou, An incremental approach for inducing knowledge from dynamic information systems, Fund. Inform. 94 (2) (2009) 245–260.
[36] X.B. Yang, J.Y. Yang, C. Wu, D.J. Yu, Dominance-based rough set approach and knowledge reductions in incomplete ordered information system, Inform. Sci. 178 (4) (2008) 1219–1234.
[37] X.B. Yang, Z.H. Chen, H.L. Dou, M. Zhang, J.Y. Yang, Neighborhood system based rough set: Models and attribute reductions, Int. J. Uncertainty Fuzz. Knowl.-Based Syst. 20 (3) (2012) 399–419.
[38] C.Z. Wang, Y. Huang, M.W. Shao, X.D. Fan, Fuzzy rough set-based attribute reduction using distance measures, Knowl.-Based Syst. 164 (2019) 205–212.
[39] C.Z. Wang, Y. Huang, M.W. Shao, D.G. Chen, Uncertainty measures for general fuzzy relations, Fuzzy Sets and Systems 360 (2019) 82–96.
[40] D. Ślęzak, A. Janusz, Ensembles of bireducts: towards robust classification and simple representation, in: International Conference on Future Generation Information Technology, Springer, 2011, pp. 64–77.
[41] D. Ślęzak, M. Grzegorowski, A. Janusz, M. Kozielski, S.H. Nguyen, et al., A framework for learning and embedding multi-sensor forecasting models into a decision support system: A case study of methane concentration in coal mines, Inform. Sci. 451–452 (2018) 112–133.
[42] J.J. Bi, C.S. Zhang, An empirical comparison on state-of-the-art multi-class imbalance learning algorithms and a new diversified ensemble learning scheme, Knowl.-Based Syst. 158 (2018) 81–93.
[43] T.R. Hoens, Q. Qian, N.V. Chawla, Z.H. Zhou, Building decision trees for the multi-class imbalance problem, in: Advances in Knowledge Discovery and Data Mining, Springer, Berlin Heidelberg, 2012, pp. 122–134.
[44] C.S. Zhang, C.C. Liu, X.L. Zhang, G. Almpanidis, An up-to-date comparison of state-of-the-art classification algorithms, Expert Syst. Appl. 82 (2017) 128–150.
[45] C.S. Zhang, J.J. Bi, S.X. Xu, E. Ramentol, G.J. Fan, et al., Multi-imbalance: An open-source software for multi-class imbalance learning, Knowl.-Based Syst. 174 (2019) 137–143.
[46] X.F. Deng, Y.Y. Yao, An information-theoretic interpretation of thresholds in probabilistic rough sets, in: International Conference on Rough Sets and Knowledge Technology, Springer, 2012, pp. 369–378.
[47] M.K. Sayadi, M. Heydari, K. Shahanaghi, Extension of VIKOR method for decision making problem with interval numbers, Appl. Math. Model. 33 (2009) 2257–2262.
[48] Z.S. Xu, Q.L. Da, The uncertain OWA operator, Int. J. Intell. Syst. 17 (2002) 569–575.
[49] W. Pedrycz, M.L. Song, Analytic hierarchy process (AHP) in group decision making and its optimization with an allocation of information granularity, IEEE Trans. Fuzzy Syst. 19 (3) (2011) 527–539.
[50] Z.H. Zhou, Machine Learning, Tsinghua University Press, Beijing, 2017.