Omega 111 (2022) 102642

Contents lists available at ScienceDirect

Omega

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

Data-driven method to learning personalized individual semantics to support linguistic multi-attribute decision making ☆

Cong-Cong Li [a,b], Yucheng Dong [c,d,*], Haiming Liang [c], Witold Pedrycz [e,f,g], Francisco Herrera [h,i]

[a] School of Economics and Management, Southwest Jiaotong University, Chengdu, China
[b] Key Laboratory of Service Science and Innovation of Sichuan Province, Southwest Jiaotong University, Chengdu, China
[c] Center for Network Big Data and Decision-Making, Business School, Sichuan University, Chengdu, China
[d] School of Management, Shenzhen Institute of Information Technology, Shenzhen, China
[e] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada
[f] Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
[g] Faculty of Engineering, King Abdulaziz University, Jeddah, Saudi Arabia
[h] Andalusian Research Institute on Data Science and Computational Intelligence (DaSCI), University of Granada, Granada, Spain
[i] Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 26 July 2021
Accepted 25 March 2022
Available online 29 March 2022

Keywords:
Decision analysis
Multi-attribute decision making
Personalized individual semantics
Learning function
Classification

ABSTRACT

In parallel with the development of information and network technology, large amounts of data are being generated by the Internet, and data-driven methodologies are now often being used in decision-making. Recent studies have investigated personalized individual semantics (PIS) in various decision-making contexts to model a fact that words mean different things to different people. However, few studies have investigated PIS in the context of multi-attribute decision-making (MADM). In MADM, in addition to multi-attribute linguistic information, pre-existing classification of the alternatives is always present, which have not been considered in prior research. Most previous studies have simply demonstrated the feasibility of PIS methods with numerical examples using small-scale models, and not with realistic datasets. Therefore, in this study, we propose a data-driven learning model to analyze the PIS of decision makers to support a multi-attribute decision-making model that considers pre-existing classification of the alternatives. Specifically, we first propose a PIS multi-attribute learning function to define a general computation form for comprehensive evaluation of the value of alternatives. Then, considering this pre-existing classification of the alternatives, a PIS learning model is constructed by analyzing the relations between calculated values of alternatives and corresponding class assignments to obtain personalized numerical scales of linguistic terms for a decision maker. Finally, we present a case study based on two datasets and a comparison with other methods to justify the feasibility of the proposed model.

1. Introduction

Multi-attribute decision-making (MADM) involves an individual or a group of decision makers selecting an option by evaluating a set of alternatives according to multiple attributes [1,2,3,4]. In realistic MADM problems, decision makers often prefer to use language and linguistic terms to express their preferences for evaluating the objects or alternatives, and such information is conventionally included in multi-attribute linguistic decision matrices to represent their preferences [5].

This means there is a need for computing with words (CWW) [6,7] when dealing with linguistic preferences in decision making. Obviously, words mean different things to different people [8,9], and type-2 fuzzy sets [8] are commonly used to deal with this issue in CWW. The CWW model of type-2 fuzzy sets is a useful tool to deal with multiple meanings, but this model does not represent specific semantics for individuals. For example, if a family with three members wants to buy a car, perhaps the members of the family all think the car is "good". However, the numerical decision-making meaning of the word "good" might be accurately represented as 0.9 for one member, and 0.7 for the other two members. This reflects PIS among the three members. Porro et al. [10] proposed the concept of perceptual maps to represent the differences between the decision makers' semantics of linguistic terms in MADM. Recently, Li et al. [11] presented a personalized individual semantics (PIS) model to customize individual nu-

☆ Area: Decision Analysis and Preference-Driven Analytics. This manuscript was processed by Associate Editor Salvatore Corrente.

* Corresponding author.
E-mail address: ycdong@scu.edu.cn (Y. Dong).

https://doi.org/10.1016/j.omega.2022.102642

merical semantics by means of a numerical scale model [12,13] and a 2-tuple linguistic model [14]. Based on the PIS model in [11], Huang and Li [15] and Li et al. [16] presented a consensus decision-making model with personalized interval numerical scales representing linguistic preferences. Tang et al. [17,18] and Li et al. [19] proposed consistency-based goal programming models to obtain personalized numerical scales (PNS) that describe distributed linguistic representations. Zhang et al. [20] proposed an optimization-based PIS model with comparative linguistic expression preferences, taking into account individual self-confidence. In addition, PIS models have also been applied in failure mode and effects analysis [21] and in research on opinion dynamics [22].

In PIS learning, the essential task is to customize the PIS of decision makers via their PNS. There are avenues for further exploration:

(1) Previous studies regarding PIS learning were conducted in the context of group decision-making models with linguistic preference relations, and they provide consistency-driven optimization models to represent PIS among decision makers by obtaining their PNS. In realistic applications, many decision-making problems occur in linguistic MADM (LMADM) contexts [23], but only a few studies have examined PIS in the context of LMADM.

(2) There are insufficiently explored data-driven LMADM (DD-LMADM) problems, which usually relate to the multi-attribute decision matrix and pre-existing classification of the alternatives. In DD-LMADM, the pre-existing classification of the alternatives stands for a classification of alternatives (or objects). In the method proposed in the paper, we assume that a classification of the alternatives exists in addition to the description using the attributes. For example, with the development of the tourism market, tourists seeking to go on a trip usually select their hotels on tourism websites, such as Ctrip, by evaluating multiple attributes of the hotels, including price, location, services, ambience and so on. Additionally, the hotels listed on tourism websites may be divided into classes according to the quality of hotels (e.g., five stars; four stars; three stars and two stars). However, most studies have only investigated LMADM based on data on multi-attribute linguistic information, without studying the pre-existing classification on alternative used in DD-LMADM.

To overcome these limitations, in this study, we develop a data-driven method for learning of PIS to support LMADM based on both multi-attribute linguistic information and pre-existing classification of the alternatives. Preference learning [24,25,26,27,28] is an important field in machine learning, and it deals with the learning of preferences in datasets by constructing a model from a given training sample. Inspired by the ideas of PIS and preference learning, our study includes the following stages:

Constructing a PIS learning model. A PIS multi-attribute learning function is first proposed to define a general computation form for evaluating the comprehensive values of alternatives. Then, based on the proposed function, we construct a PIS learning model for DD-LMADM to calculate the PNS of linguistic terms for decision makers, by constraining the value difference of pairs of alternatives from consecutive classes and the same class.

Application of the PIS learning model. We use real datasets to test our PIS learning model. A comparison with existing methods that do not implement PIS is included. Our results show that the integration of PIS learning in DD-LMADM systems may improve consistency between linguistic preferences over attributes and their corresponding class assignments.

The personalization of linguistic preferences and the integration of the pre-existing classification of the alternatives in LMADM show the advantages of the proposed method for dealing with the individual linguistic preference understanding and in improving the consistency of decision results.

The study is arranged as follows. Section 2 introduces the necessary preliminaries to develop our proposal. Section 3 presents a description for PIS learning in DD-LMADM. Section 4 details a PIS multi-attribute learning function, based on which our data-driven PIS learning model is proposed. Section 5 includes case studies based on two learning datasets to illustrate our proposed model in a real LMADM context. Section 6 provides a comparison with existing approaches that do not implement PIS. Finally, Section 7 concludes the paper.

2. Preliminaries

In this section, we introduce basic knowledge regarding linguistic approaches based on membership functions, a 2-tuple linguistic model, and PIS based on numerical scale.

2.1. Linguistic approach based on membership function

The concept and applications of linguisitc variables are introduced in [6]. Let $S = \{s_0, s_1, ..., s_g\}$ be a linguistic term set. The linguistic term s_i represents a possible value for a linguistic variable where $s_i > s_j$ if and only if $i > j$.

Triangular membership functions [14,29,6] are one of the most commonly used membership functions in fuzzy sets to represent the semantics of linguistic terms, described:

$$A(x)=\begin{cases} 0, x < a \\ (x-a)/(b-a), a \le x < b \\ (c-x)/(c-b), b \le x < c \\ 0, c \le x \end{cases} \tag{1}$$

where a, b and c are parameters of $A(x)$.

These triangular fuzzy numbers can be used as appropriate descriptors to represent the semantics of linguistic terms. For example, we can represent a linguistic term set S with seven terms as follows:

$$S = \{s_0 : N, s_1 : VL, s_2 : L, s_3 : M, s_4 : H, s_5 : VH, s_6 : P\}$$

Using the labels based on triangular membership function, the semantics of linguistic terms in S can be assigned as follows:

$$A(N) = (0, 0, 0.17), A(VL) = (0, 0.17, 0.33), A(L) = (0.17, 0.33, 0.5),$$
$$A(M) = (0.33, 0.5, 0.67), A(H) = (0.5, 0.67, 0.83),$$
$$A(VH) = (0.67, 0.83, 1), A(P) = (0.83, 1, 1)$$

Then, a CWW model can be constructed based on triangular membership function semantics of linguistic terms. Some basic operations on triangular membership functions can be found in [30,6].

2.2. 2-tuple linguistic model and PIS based on numerical scale

Herrera and Martínez [14] argued that there is information loss in CWW based on membership functions, and thus, they proposed a linguistic representation model based on symbolic translation, called the 2-tuple linguistic model, for CWW with linguistic 2-tuples.

Definition 1 [14]. Let $S = \{s_0, s_1, ..., s_g\}$ be a linguistic term set and $\beta \in [0, g]$ be a value representing the result of a symbolic aggregation operation. The transformation functions between 2-tuples and numerical values are defined as:

$$\Delta : [0, g] \to \overline{S} \tag{2}$$

being

$\Delta(\beta) = (s_i, \alpha)$ with

$$\begin{cases} s_i, i = round(\beta) \\ a = \beta - i, a \in [-0.5, 0.5) \end{cases} \tag{3}$$

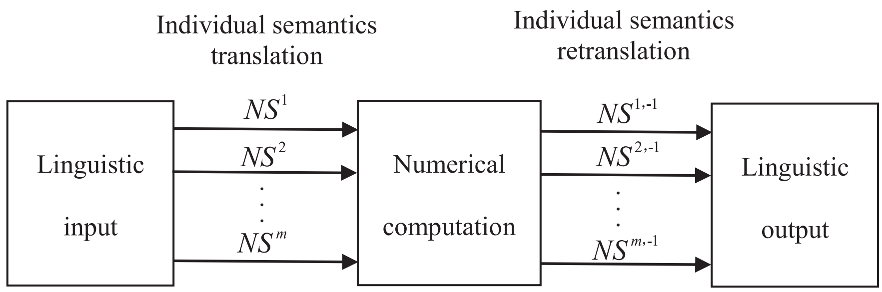

Fig. 1. The framework for the linguistic model with PIS [35].

The inverse function of Δ, $\Delta^{-1}:\bar{S}\rightarrow[0,g]$ is defined as $\Delta^{-1}(s_i,\alpha)=i+\alpha$.

The 2-tuple linguistic model provides a popular tool for linguistic decision making, but it only deals with linguistic term sets that are uniformly and symmetrically distributed. i.e., balanced linguistic term sets.

Dong et al. [12] developed a numerical scale model as an extension of the 2-tuple linguistic model to deal with both balanced and unbalanced linguistic term sets.

Definition 2 [12]. Let $S=\{s_0,s_1,\ldots,s_g\}$ be a set of linguistic terms, and R be the set of real numbers. A function $NS:S\rightarrow R$ is called a numerical scale of S, and $NS(s_i)$ is the numerical value of NS.

Definition 3 [12]. The numerical scale NS for (s_i,α) is defined as follows:

$$NS(s_i,\alpha)=\begin{cases}NS(s_i)+\alpha\times(NS(s_{i+1})-NS(s_i)),\alpha\geq 0\\ NS(s_i)+\alpha\times(NS(s_i)-NS(s_{i-1})),\alpha<0\end{cases} \tag{4}$$

If $NS(s_i)<NS(s_{i+1})$, for $i=0,1,\ldots,g-1$, then the NS on S is ordered.

The inverse operator of the numerical scale NS is defined as

$$NS^{-1}:R\rightarrow\bar{S} \tag{5}$$

with

$$NS^{-1}(r)=\begin{cases}\left(s_i,\dfrac{r-NS(s_i)}{NS(s_{i+1})-NS(s_i)}\right),NS(s_i)<r<\dfrac{NS(s_i)+NS(s_{i+1})}{2}\\ \left(s_i,\dfrac{r-NS(s_i)}{NS(s_i)-NS(s_{i-1})}\right),\frac{NS(s_{i-1})+NS(s_i)}{2}\leq r\leq NS(s_i)\end{cases} \tag{6}$$

Numerical scale models provide a uniform framework [31] to connect the various 2-tuple linguistic models. In [31], it shows the setting of $NS(s_i)=\Delta^{-1}(s_i)$ yields the 2-tuple linguistic model.

To represent different understandings of words by decision makers, Li et al. [11] presented a framework with three processes to handle linguistic information in linguistic decision-making with PIS (Graphically, see Fig. 1). Specifically, it translates linguistic terms into numerical values, which would be used in a numerical computation process representing individual semantics. Then, the framework retranslates the output of these numerical scales into linguistic values that are easier for decision makers to understand.

In Fig. 1, NS^k is a numerical scale on S associated with decision maker $e_k(k=1,2,\ldots,m)$, and the value of $PVV^k=\{PVV_1^k,PVV_2^k,\ldots,PVV_n^k\}$ represents the PIS of decision maker e_k associated with the term $s_i(i=0,1,\ldots,g)$. Furthermore, to represent the specific semantics of decision makers, Li et al. [11] proposed consistency-driven optimization models by calculating the PNS of linguistic terms.

3. Data-driven linguistic multi-attribute decision making problem

In this section, we formulate a DD-LMADM system to study PIS learning. In our DD-LMADM model, there are two kinds of data:

(1) The multi-attribute linguistic data. Considering the complexity of decision-making environments and the diversity of linguistic evaluations, the linguistic term sets used to evaluate attributes may be multi-granular [32,33], that is, linguistic term sets of varying cardinality and/or semantics could be used to express decision makers' opinions on the set of alternatives.

(2) The pre-existing classification of the alternatives. In DD-LMADM, the alternatives are often classified into several predefined classes according to performance. For example, when buying cars, the customers would evaluate the cars with linguistic preferences on multiple attributes, and meanwhile, there is the cars' classification information: economic car and luxury car.

Then, we describe the following notations that will be used in this system.

(1) Multi-attribute linguistic data:

(i) $E=\{e_1,e_2,\ldots,e_m\}$: The set of decision makers;

(ii) $R=\{r_1,r_2,\ldots,r_q\}(q\geq 2)$: The set of attributes for the decision makers in E to evaluate the alternatives;

(iii) $S^j=\{s_0^j,s_1^j,\ldots,s_{gj}^j\}$: The linguistic term set, associated with the attribute r_j, to be used for the decision makers in E to express their linguistic assessments;

(iv) $X^k=\{x_i^k|i=1,2,\ldots,n_k\}$: The set of alternatives, associated with e_k;

(v) $L^k=(l_{ij}^k)_{n_k\times q}$: The multi-attribute linguistic decision matrix provided by e_k, where l_{ij}^k indicates the linguistic assessment of alternative x_i^k on attribute r_j, associated with e_k.

(2) Pre-existing classification of the alternatives:

(i) $H=\{1,2,\ldots,h\}$ is the set of classification labels;

(ii) $c(x_i^k)$ is the label of x_i^k, where $c(x_i^k)\in H$; and the smaller value of $c(x_i^k)$ indicate the better alternative x_i^k;

(iii) $C^k=\{C_1^k,C_2^k,\ldots,C_h^k\}$ is the classifications over X^k, where $C_v^k=\{x_i^k|c(x_i^k)=v\}(v\in H)$. Clearly $\bigcup_{v=1}^{h}C_v^k=X^k$ and $C_v^k\cap C_e^k=\emptyset$ $(v\neq e)$. We call C^k the pre-existing classification of the alternatives in X^k, associated with e_k.

In CWW, the process of assigning PNS to linguistic terms models the PIS of linguistic terms used by decision makers. In this DD-LMADM system, although we have linguistic assessments on multiple attributes and information classifying the alternatives, we do not know specific PIS for decision makers. Thus, the main objective of this study is to propose a data-driven method for learning of PIS to support LMADM.

The data-driven method for calculating PNS is implemented with the following two steps:

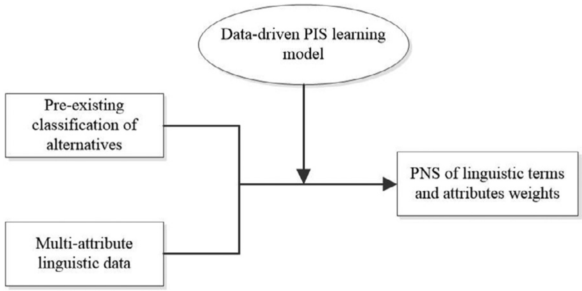

Fig. 2. Framework of learning PIS in the DD-LMADM.

Define the PIS multi-attribute learning function. We present this function to define a general computation form to comprehensively evaluate the alternatives.

Construct the data-driven PIS learning model. Based on this PIS multi-attribute learning function, we construct a data-driven optimization model based on multi-attribute linguistic data and pre-existing classification data to obtain the PNS of linguistic terms for decision makers.

Fig. 2 provides the resolution framework of the proposed model. In Fig. 2, based on two kinds of data (i.e., multi-attribute linguistic data and pre-existing classification of alternatives), we obtain the PNS through solving the data-driven PIS learning model. The values of PNS among decision makers are different, which shows the PIS of decision makers.

Notably, we consider two cases for the attributes' weights: unknown attribute weights and pre-established attribute weights.

In DD-LMADM, it is assumed that decision makers express their linguistic assessments through simple linguistic terms. However, it is possible that decision makers use complex linguistic representation structures to express their preferences. In this situation, flexible linguistic expressions [34] can provide a general tool to model decision makers' linguistic assessments.

4. A PIS learning method in the DD-LMADM

In this section, we propose a new method for calculating PIS in a DD-LMADM model. Specifically, we first propose a PIS multi-attribute learning function to comprehensively evaluate the values of alternatives. Then, based on the PIS multi-attribute learning function, we construct a PIS learning model with the multi-attribute linguistic data and pre-existing classification of the alternatives to compute PNS in the DD-LMADM model system.

4.1. PIS multi-attribute learning function

Let $S^j=\{s_0^j,s_1^j,...,s_{gj}^j\}$ be the linguistic term set associated with the attribute r_j. To study PIS in the DD-LMADM formal system, we introduce the PNS, $PNS^k(l_{ij}^k)$, to define the individual numerical meaning of l_{ij}^k over S^j, which is used to characterize PIS of decision makers. Using this definition and based on the multi-attribute additive model [35], we define the personalized comprehensive value for alternative x_i associated with e_k as follows:

$$F^k(x_i)=\sum_{j=1}^{q} w_j^k\times PNS^k(l_{ij}^k) \tag{7}$$

where $W^k=(w_1^k,w_2^k,...,w_q^k)^T$ is the weight vector of the attributes associated with decision maker e_k, which satisfies the requirements $w_j^k\geq 0$ and $\sum_{j=1}^{q} w_j^k=1$.

Notably, in accordance with the characteristics of realistic decision making, the preferences provided by decision makers on attributes can be divided into two types: qualitative and quantitative. If the preference l_{ij}^k is quantitative, we assume that l_{ij}^k is bounded to the interval [0,1] and set $PNS^k(l_{ij}^k)=l_{ij}^k$ in this study. If the preference l_{ij}^k is qualitative, without any loss of generality, the range of $PNS^k(l_{ij}^k)$ is expressed in the form:

$$PNS^k(l_{ij}^k)\begin{cases}=0,\ if\, l_{ij}^k=s_0^j\\ \in(0,1),\ for\, l_{ij}^k\neq s_0^j and\, l_{ij}^k\neq s_{gj}^j\\ =1,\ if\, l_{ij}^k=s_{gj}^j\end{cases} \tag{8}$$

Eq. (8) provides the range of personalized numerical scales. Specifically, if $l_{ij}^k=s_0^j$, we set the numerical scale of l_{ij}^k is $PNS^k(l_{ij}^k)$=0; if $l_{ij}^k=s_{gj}^j$, we set the numerical scale of l_{ij}^k is $PNS^k(l_{ij}^k)$= 1; for $l_{ij}^k\neq s_0^j$ and $l_{ij}^k\neq s_{gj}^j$, we set $PNS^k(l_{ij}^k)\in(0,1)$. It is worth pointing out that the range of the $PNS^k(l_{ij}^k)$ may be different for different decision-making problems, which can be determined by the demand of decision makers.

Then, in order to calculate the values of $PNS^k(l_{ij}^k)$, we introduce a deviation variable $\varepsilon_t^{j,k}=PNS^k(s_t^j)-PNS^k(s_{t-1}^j)$, which denotes the deviation between the PNS of linguistic terms s_t^j and s_{t-1}^j $(t=1,2,...,gj)$. Clearly, $\varepsilon_t^{j,k}>0$ and $\sum_{t=1}^{gj}\varepsilon_t^{j,k}=1$.

Without loss of generality, we assume $l_{ij}^k=s_\gamma^j\,(\gamma\in\{0,1,...,gj\})$. Then, $PNS^k(l_{ij}^k)$ is formulated as follows:

$$PNS^k(l_{ij}^k)=PNS^k(s_\gamma^j)=\sum_{t=1}^{\gamma}\varepsilon_t^{j,k} \tag{9}$$

Next, let $\varepsilon^{j,k}=(\varepsilon_1^{j,k},\varepsilon_2^{j,k},...,\varepsilon_{gj}^{j,k})^T$ and $O^j(l_{ij}^k)=(o_1^j(l_{ij}^k),o_2^j(l_{ij}^k),...,o_{gj}^j(l_{ij}^k))^T$, where

$$o_t^j\left(l_{ij}^k\right)=o_t^j\left(s_\gamma^j\right)=\begin{cases}1,\ if\gamma\geq t\\ 0,\ if\gamma<t\end{cases} \tag{10}$$

Eq. (9) can be equivalently transformed into the following form represented by the vectors $\varepsilon^{j,k}$ and $O^j(l_{ij}^k)$:

$$PNS^k\left(l_{ij}^k\right)=\left(\varepsilon^{j,k}\right)^T O^j\left(l_{ij}^k\right) \tag{11}$$

In the following, we list an example with a linguistic term set $S^j=\{s_0^j,...,s_6^j\}$ to show the use of $\varepsilon_t^j(t=1,2,...,6)$ to denote the PNS of linguistic terms, as shown in Fig. 3.

Furthermore, based on Eq. (7), the comprehensive value $F^k(x_i)$ of alternative x_i associated with decision maker e_k is rewritten as follows:

$$F^k(x_i)=\sum_{j=1}^{q} w_j^k\left(\varepsilon^{j,k}\right)^T O^j\left(l_{ij}^k\right)=\left(W^k\right)^T U^k(x_i) \tag{12}$$

where $W^k=(w_1^k,w_2^k,...,w_q^k)^T$ and $U^k(x_i)=((\varepsilon^{1,k})^T O^1(l_{i1}^k),(\varepsilon^{2,k})^T O^2(l_{i2}^k),...,(\varepsilon^{q,k})^T O^q(l_{iq}^k))^T$.

Because the use of Eq. (12) can evaluate the comprehensive value of linguistic preferences in this DD-LMADM context, as well as compare the alternatives in different classes, thus we use Eq. (12) as our PIS multi-attribute learning function.

4.2. A data-driven PIS learning model in DD-LMADM

As defined earlier, $C^k=\{C_1^k,...,C_h^k\}$ with $C_i^k\succ C_{i+1}^k$ is a set of classes of alternatives, which indicates that the comprehensive values of the alternatives in class C_i^k should be no worse than those in class C_{i+1}^k, i.e., $F^k(x_a)\geq F^k(x_b)$ for $\forall x_a\in C_i^k$, $\forall x_b\in C_{i+1}^k$ and $k=1,2,...,m$.

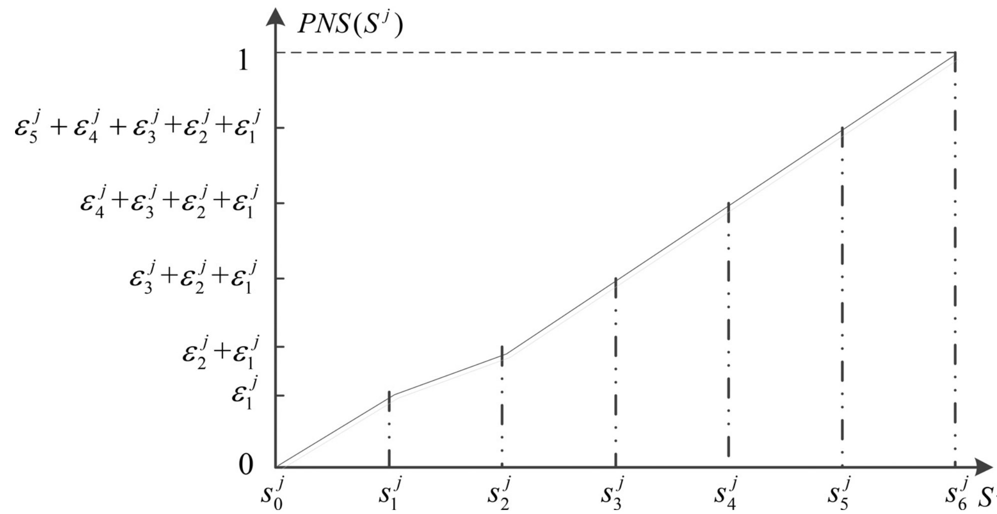

Fig. 3. Representation of the PNS values of s_t^j with $\varepsilon_t^j (t = 1, 2, ..., 6)$.

In this data-driven PIS learning model, inspired by the idea of preference learning [27], we hope to achieve two objectives:

Maximize the minimum average comprehensive value difference among alternatives between any two consecutive classes, and

Minimize the comprehensive value difference among alternatives from the same class.

Let d^k ($d^k \geq 0$) be the minimum value difference for the average comprehensive values between any two consecutive classes associated with decision maker e_k, thus

$$\frac{\sum_{x_a \in C_i^k} F^k(x_a)}{\#C_i^k} - \frac{\sum_{x_b \in C_{i+1}^k} F^k(x_b)}{\#C_{i+1}^k} = \frac{\sum_{x_a \in C_i^k} \left(\mathrm{W}^k\right)^{\mathrm{T}} \mathrm{U}^k(x_a)}{\#C_i^k} - \frac{\sum_{x_b \in C_{i+1}^k} \left(\mathrm{W}^k\right)^{\mathrm{T}} \mathrm{U}^k(x_b)}{\#C_{i+1}^k} \geq d^k \text{for} i = 1, 2, ..., h-1 \tag{13}$$

where $\#C_i^k$ is the cardinality of $C_i^k (i = 1, 2, ..., h)$.

Then, the first objective (i) is represented as follows:

$$\min -d^k \tag{14}$$

Next, let $d^{k'}$ denote the value difference of the comprehensive values among alternatives from a class C_i^k $(i = 1, 2, ..., h)$ associated with e_k. To achieve the second objective (ii), we use

$$\min d^{k'} \tag{15}$$

where

$$d^{k'} = \sum_{i=1}^{p} \sum_{x_a, x_b \in C_i} \left(\left(\mathrm{W}^k\right)^{\mathrm{T}} \mathrm{U}^k(x_a) - \left(\mathrm{W}^k\right)^{\mathrm{T}} \mathrm{U}^k(x_b) \right)^2 = \left(\mathrm{W}^k\right)^{\mathrm{T}} \mathrm{Q}^k \left(\mathrm{W}^k\right) \tag{16}$$

with $\mathrm{Q}^k = \sum_{i=1}^{p} \sum_{x_a, x_b \in C_i} (\mathrm{U}^k(x_a) - \mathrm{U}^k(x_b))(\mathrm{U}^k(x_a) - \mathrm{U}^k(x_b))^{\mathrm{T}}$.

We use the method based on membership function [36] to combine the above two objectives (see, Eqs. (14) and (15)), we obtain the objective function of the proposed model as follows,

$$\max \alpha \lambda_1^k + \beta \lambda_2^k \tag{17}$$

where λ_1^k and λ_2^k denote the coefficient of the objectives (i) and (ii):

$$\lambda_1^k = \frac{\max\left(-d^k\right) - \left(-d^k\right)}{\max\left(-d^k\right) - \min\left(-d^k\right)} \text{and} \lambda_2^k = \frac{\max\left(d^{k'}\right) - d^{k'}}{\max\left(d^{k'}\right) - \min\left(d^{k'}\right)}.$$

In Eq. (17) α and β denote the weights of λ_1^k and λ_2^k, respectively, satisfying $\alpha, \beta \in [0, 1]$ and $\alpha + \beta = 1$. A larger value of α (or β) demonstrates a greater importance on objective (i) (or (ii)). In this study, we assume that objectives (i) and (ii) have equal importance, and thus set $\alpha = \beta = 0.5$. Notably, the values of α and β should be determined according to practical decision problems, and our model is general and applicable when setting different values of α and β.

The values of $\max(-d^k)$, $\min(-d^k)$, $\min(d^{k'})$ and $\max(d^{k'})$ can be determined respectively by substituting the objective functions in the following PIS learning models (P_1 and P_2).

We consider two cases to construct the PIS learning model:

Case A. PIS learning with unknown attribute weights. Based on the above, we construct our PIS learning model in a DD-LMADM system with unknown attribute weights as follows:

$$\begin{cases} \max \alpha \lambda_1^k + \beta \lambda_2^k \\ s.t. \begin{cases} \lambda_1^k = \frac{\max\left(-d^k\right) - \left(-d^k\right)}{\max\left(-d^k\right) - \min\left(-d^k\right)} \\ \lambda_2^k = \frac{\max\left(d^{k'}\right) - d^{k'}}{\max\left(d^{k'}\right) - \min\left(d^{k'}\right)} \\ \frac{\sum\limits_{x_a \in C_i^k} \left(\mathrm{W}^k\right)^{\mathrm{T}} \mathrm{U}^k(x_a)}{\#C_i^k} - \frac{\sum\limits_{x_b \in C_{i+1}^k} \left(\mathrm{W}^k\right)^{\mathrm{T}} \mathrm{U}^k(x_b)}{\#C_{i+1}^k} \geq d^k, i = 1, 2, ..., h-1 \\ d^{k'} = \left(\mathrm{W}^k\right)^{\mathrm{T}} \mathrm{Q}^k \mathrm{W}^k \\ \varepsilon_t^{j,k} > 0, t = 1, 2, ..., gj; j = 1, 2, ..., q \\ \sum\limits_{t=1}^{gj} \varepsilon_t^{j,k} = 1, j = 1, 2, ..., q \\ 0 < w_j^k < 1, j = 1, 2, ..., q \\ \sum\limits_{j=1}^{q} w_j^k = 1 \end{cases} \end{cases}$$

where $w_j^k (j = 1, ..., q)$ and $\varepsilon_t^{j,k} (t = 1, 2, ..., gj; j = 1, 2, ..., q)$ are decision variables. This model is denoted as model P_1.

Case B. PIS learning with pre-established attribute weights. We construct the PIS learning model in a DD-LMADM system with pre-

Table 1
Summary of dataset of car information.

Dataset	Decision makers	#Alternatives	#Attributes	#Classes	Distribution of classes
Car Evaluation	e_1	138	6	4	55/58/10/15
	e_2	138	6	4	53/51/20/14
	e_3	125	6	4	51/45/15/14
	e_4	129	6	4	54/47/15/13

established attribute weights as follows:

$$\begin{cases} \max \alpha\lambda_1^k+\beta\lambda_2^k \\ \text{s.t.} \begin{cases} \lambda_1^k=\frac{\max(-d^k)-(-d^k)}{\max(-d^k)-\min(-d^k)} \\ \lambda_2^k=\frac{\max\left(d^{k'}\right)-d^{k'}}{\max\left(d^{k'}\right)-\min\left(d^{k'}\right)} \\ \frac{\sum_{x_a\in C_i^k}(W^k)^{\mathrm{T}}U^k(x_a)}{\#C_i^k}-\frac{\sum_{x_b\in C_{i+1}^k}(W^k)^{\mathrm{T}}U^k(x_b)}{\#C_{i+1}^k}\geq d^k, i=1,2,...,h-1 \\ d^{k'}=\left(W^k\right)^{\mathrm{T}}Q^kW^k \\ \varepsilon_t^{j,k}>0, t=1,2,...,gj;\ j=1,2,...,q \\ \sum_{t=1}^{gj}\varepsilon_t^{j,k}=1, j=1,2,...,q \end{cases} \end{cases}$$

where $\varepsilon_t^{j,k}(t=1,2,...,gj;\ j=1,2,...,q)$ are decision variables. This model is denoted as model P_2.

Since models P_1 and P_2 belong to a type of quadratic programming model, we can directly use some software packages (e.g., Matlab, Cplex) to solve models P_1 and P_2. In models P_1 and P_2, based on the representation $\varepsilon_t^{j,k}=PNS^k(s_t^j)-PNS^k(s_{t-1}^j)$ $(t=1,2,...,gj;\ j=1,2,...,q)$, we can learn the PNS of linguistic terms associated with attributes r_j for decision makers e_k, i.e., $\{PNS^k(s_0^j), PNS^k(s_1^j),...,PNS^k(s_{gj}^j)\}$, which reflect the PIS among decision makers. The difference between models P_1 and P_2 is that we learn the attribute weights from P_1 and in P_2 the attribute weights are pre-established. Moreover, in the case of attribute weights with partial information, we can add constraints on partial weight information to P_1 to support PIS learning.

Furthermore, based on the PNS obtained from models P_1 or P_2, the multi-attribute linguistic decision matrix $L^k=(l_{ij}^k)_{n\times q}$ can be converted into a multi-attribute numerical decision matrix $V^k=(v_{ij}^k)_{n\times q}$, i.e., $v_{ij}^k=PNS^k(l_{ij}^k)$, which can be further applied in the selection process for alternative rankings based on the classical multi-attribute numerical decision making methods.

5. Case study

In this section, we illustrate the data-driven PIS learning method with two real datasets: a car evaluation dataset and a house evaluation dataset.

5.1. Car evaluation dataset

The car evaluation database used here was derived from a simple hierarchical decision model originally developed for the demonstration of DEX [37]. The cars were evaluated linguistically based on six attributes $\{r_1,r_2,...,r_6\}$. Specifically, r_1 is the buying price, r_2 is the price of the maintenance, r_3 is the number of doors, r_4 is the capacity in terms of passengers, r_5 is the size of luggage space, and r_6 is the estimated safety of the car. In order to illustrate the proposed model, in this case study we classify the whole dataset into four subsets randomly, each subset containing 138, 138, 125, 129 alternatives, respectively. We assume that these subsets are provided by four decision makers $\{e_1,e_2,e_3,e_4\}$, and each decision maker has his/her PIS. In addition, in each subset the cars are categorized into four classes based on their performances, denoted as C_1^k, C_2^k, C_3^k, and C_4^k $(k=1,2,3,4)$. For example, in the subset associated with e_1, there are 55 alternatives belonging to class C_1^1, 58 alternatives belonging to C_2^1, 10 alternatives belonging to C_3^1, and 15 alternatives belonging to C_4^1. A detailed information about the car evaluation dataset is provided in Table 1.

The scales $S^j(j=1,2,...,6)$ are used to evaluate the attributes r_j, and they are listed as follows:

$$S^1=\left\{s_0^1=verylow, s_1^1=low, s_2^1=medium, s_3^1=high, s_4^1=veryhigh\right\}$$

$$S^2=\left\{s_0^2=verylow, s_1^2=low, s_2^2=medium, s_3^2=high, s_4^2=veryhigh\right\}$$

$$S^3=\left\{s_0^3=two, s_1^3=three, s_2^3=four, s_3^3=five, s_4^3=morethanfive\right\}$$

$$S^4=\left\{s_0^4=two, s_1^4=four, s_2^4=morethanfour\right\}$$

$$S^5=\{s_0^5=small, s_1^5=medium, s_2^5=big\}$$

$$S^6=\{s_0^6=low, s_1^6=medium, s_2^6=high\}.$$

In this study, we consider the same weight for the two objectives Eqs. (14) and ((15)) presented in Model P_1 and set $\alpha=\beta=0.5$. Based on Model P_1, we construct the data-driven PIS learning model for the learning datasets of car evaluation as follows,

$$\begin{cases} \max 0.5\lambda_1^k+0.5\lambda_2^k \\ s.t. \\ \frac{\sum_{x_a\in C_i^k}(W^k)^{\mathrm{T}}U^k(x_a)}{\#C_i^k}-\frac{\sum_{x_b\in C_{i+1}^k}(W^k)^{\mathrm{T}}U^k(x_b)}{\#C_{i+1}^k}\geq d^k, i=1,2,3 \\ d^{k'}=\left(W^k\right)^{\mathrm{T}}Q^kW^k \\ \varepsilon_t^{j,k}>0, t=1,2,3,4;\ j=1,2,3 \\ \varepsilon_t^{j,k}>0, t=1,2;\ j=4,5,6 \\ \sum_{t=1}^{4}\varepsilon_t^{j,k}=1, j=1,2,3 \\ \sum_{t=1}^{2}\varepsilon_t^{j,k}=1, j=4,5,6 \\ 0<w_j^k<1, j=1,2,...,6 \\ \sum_{j=1}^{6}w_j^k=1 \end{cases}$$

By solving the above model, we obtain the PNS of linguistic terms with respect to attributes $r_j(j=1,2,...,6)$ presented in the dataset for four decision makers $e_k(k=1,2,3,4)$ (see Tables 2–7). Fig. 4 provides the calculated PNS associated with $e_k(k=1,2,3,4)$ graphically.

Table 2
$PNS^k(s_t^1)$ with respect to attribute r_1 for $k=1,2,3,4$ and $t=0,1,2,3,4$.

	$PNS^k(s_0^1)$	$PNS^k(s_1^1)$	$PNS^k(s_2^1)$	$PNS^k(s_3^1)$	$PNS^k(s_4^1)$
$k=1$	0	0.182	0.455	0.687	1
$k=2$	0	0.381	0.582	0.699	1
$k=3$	0	0.134	0.364	0.676	1
$k=4$	0	0.276	0.576	0.778	1

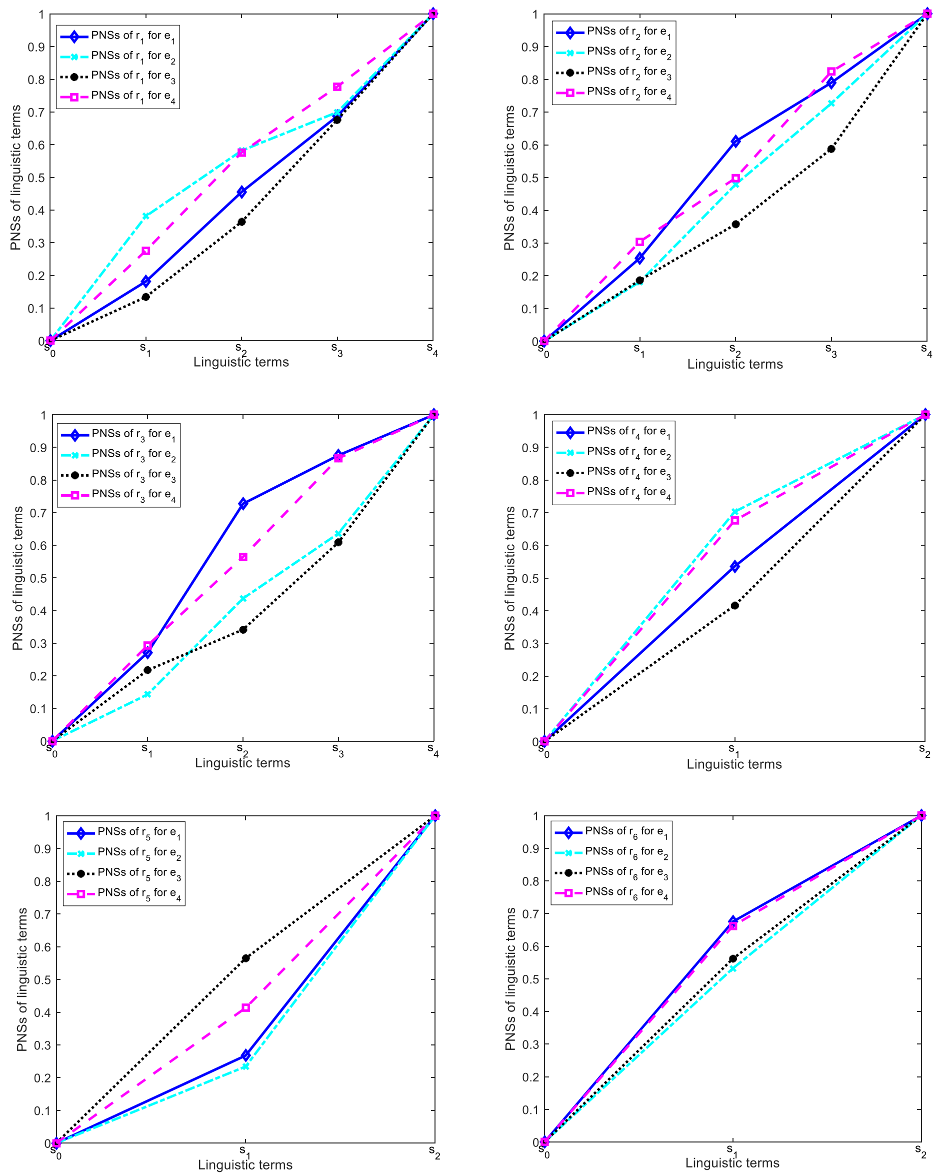

Fig. 4. PNS of linguistic terms for $e_k(k=1,2,3,4)$ with respect to $r_t(t=1,2,...,6)$.

Table 3
$PNS^k(s_t^2)$ with respect to attribute r_2 for $k=1,2,3,4$ and $t=0,1,2,3,4$.

	$PNS^k(s_0^2)$	$PNS^k(s_1^2)$	$PNS^k(s_2^2)$	$PNS^k(s_3^2)$	$PNS^k(s_4^2)$
$k=1$	0	0.254	0.611	0.79	1
$k=2$	0	0.181	0.479	0.727	1
$k=3$	0	0.186	0.358	0.588	1
$k=4$	0	0.304	0.498	0.824	1

Table 4
with respect to attribute r_3 for $k=1,2,3,4$ and $t=0,1,2,3,4$.

	$PNS^k(s_0^2)$	$PNS^k(s_1^2)$	$PNS^k(s_2^2)$	$PNS^k(s_3^2)$	$PNS^k(s_4^2)$
$k=1$	0	0.271	0.727	0.876	1
$k=2$	0	0.143	0.437	0.636	1
$k=3$	0	0.217	0.342	0.609	1
$k=4$	0	0.293	0.564	0.867	1

Table 5
$PNS^k(s_t^4)$ with respect to attribute r_4 for $k=1,2,3,4$ and $t=0,1,2$.

	$PNS^k(s_0^4)$	$PNS^k(s_1^4)$	$PNS^k(s_2^4)$
$k=1$	0	0.536	1
$k=2$	0	0.703	1
$k=3$	0	0.416	1
$k=4$	0	0.676	1

Table 6
$PNS^k(s_t^5)$ with respect to attribute r_5 for $k=1,2,3,4$ and $t=0,1,2$.

	$PNS^k(s_0^5)$	$PNS^k(s_1^5)$	$PNS^k(s_2^5)$
$k=1$	0	0.268	1
$k=2$	0	0.234	1
$k=3$	0	0.565	1
$k=4$	0	0.413	1

In addition, we also obtain the personalized attribute weights associated with each decision maker e_k for $k=1,2,3,4$ (see Table 8).

From Tables 2–7 and Fig. 4, it can be observed that the PNS of linguistic terms s_t^j with respect to attributes $r_j(j=1,2,...,6)$ are different for different decision makers, which reflects the

Table 7
$PNS^k(s_t^6)$ with respect to attribute r_6 for $k = 1, 2, 3, 4$ and $t = 0, 1, 2$.

	$PNS^k(s_0^6)$	$PNS^k(s_1^6)$	$PNS^k(s_2^6)$
$k = 1$	0	0.675	1
$k = 2$	0	0.532	1
$k = 3$	0	0.562	1
$k = 4$	0	0.662	1

Table 8
Attribute weights W^k associated with e_k for $k = 1, 2, 3, 4$.

	w_1^k	w_2^k	w_3^k	w_4^k	w_5^k	w_6^k
$k = 1$	0.16	0.237	0.201	0.076	0.128	0.198
$k = 2$	0.154	0.081	0.243	0.077	0.175	0.27
$k = 3$	0.3	0.211	0.126	0.079	0.103	0.181
$k = 4$	0.166	0.117	0.109	0.096	0.184	0.327

PIS among decision makers. We take $PNS^k(s_t^5)$ $(k = 1, 2, 3, 4; t = 0, 1, 2)$ as an example, and observe that it has $PNS^k(s_1^5)$ equal to 0.268, 0.234, 0.565, and 0.413 for decision makers e_1, e_2, e_3, and e_4, respectively. The difference of the PNS reflects the PIS. In addition, from Table 8, we can see that the computed values of attribute weights are also distinct for different decision makers.

5.2. House evaluation dataset

This case study analyzes house statistics in Den Bosch, a city with approximately 110,000 inhabitants in the Netherlands. The source of this house dataset is obtained from the reference [38]. There are eight attributes of these houses: district (r_1), area (r_2), number of bedrooms (r_3), type of the house (r_4), volume (r_5), stories (r_6), type of the garden (r_7), garage (r_8). The attributes r_2 and r_5 are quantitative, and the corresponding preferences on them are numerical values. We divide the dataset into four subsets randomly, which include 30, 30, 29, and 30 alternatives, respectively, and each one is associated with a decision maker $e_k (k \in \{1, 2, 3, 4\})$. In each subset, the alternatives (i.e., houses in Den Bosch) are divided into two classes, denoted as C_1^k and C_2^k $(k = 1, 2, 3, 4)$. Specifically, there are 14 alternatives belonging to class C_1^1 and 16 alternatives belonging to class C_2^1 for decision maker e_1; 15 alternatives belonging to class C_1^2 and 15 alternatives belonging to class C_2^2 for e_2; 14 alternatives belonging to class C_1^3 and 15 alternatives belonging to class C_2^3 for e_3; 16 alternatives belonging to class C_1^4 and 14 alternatives belonging to class C_2^4 for e_4. Detailed information about the car evaluation dataset is provided in Table 9.

The scales $S^j (j = 1, 3, 4, 6, 7, 8)$ used to evaluate the attributes $\{r_1, r_3, r_4, r_6, r_7, r_8\}$ are presented as follows:

$$S^1 = \left\{s_0^1 = \text{very}bad, s_1^1 = bad, s_2^1 = medium, s_3^1 = high, s_4^1 = veryhigh\right\}$$

$$S^3 = \{s_0^3 = \text{one}, s_1^3 = \text{two}, s_2^3 = three, s_3^3 = four, s_4^3 = five, s_5^3 = six, s_6^3 = seven\}$$

$$S^4 = \left\{s_0^4 = \text{dormitory}, s_1^4 = apartment, s_2^4 = row\ house, s_3^4 = cornerhouse, s_4^4 = semidetached\ house, s_5^4 = detached\ house, s_6^4 = villa\right\};$$

$$S^6 = \{s_0^6 = bad, s_1^6 = medium, s_2^6 = good\}$$

$$S^7 = \left\{s_0^7 = verylow, s_1^7 = low, s_2^7 = medium, s_3^7 = high, s_4^7 = veryhigh\right\}$$

$$S^8 = \{s_0^8 = no, s_1^8 = normal, s_2^8 = big\}.$$

The scales S^2 and S^5 for describing attributes r_2 and r_5 are numerical sets, and we use the numerical values presented in the dataset directly.

Consider the same weight of the two objectives Eqs. (14) and ((15)) in Model P_1 and let $\alpha = \beta = 0.5$, we input the house evaluation dataset into our data-driven PIS learning model as follows:

$$\begin{cases}
\max 0.5\lambda_1^k + 0.5\lambda_2^k \\
\text{s.t.} \\
\frac{\sum_{x_a \in C_1^k} (W^k)^T U^k(x_a)}{\# C_1^k} - \frac{\sum_{x_b \in C_2^k} (W^k)^T U^k(x_b)}{\# C_2^k} \geq d^k \\
d^{k'} = \left(W^k\right)^T Q^k W^k \\
\varepsilon_t^{j,k} > 0, t = 1, 2, 3, 4;\ j = 1, 7 \\
\varepsilon_t^{j,k} > 0, t = 1, 2, ..., 6;\ j = 3, 4 \\
\varepsilon_t^{j,k} > 0, t = 1, 2;\ j = 6, 8 \\
\sum_{t=1}^4 \varepsilon_t^{j,k} = 1, j = 1, 7 \\
\sum_{t=1}^6 \varepsilon_t^{j,k} = 1, j = 3, 4 \\
\sum_{t=1}^2 \varepsilon_t^{j,k} = 1, j = 6, 8 \\
0 < w_i^k < 1, i = 1, 2, ..., 8 \\
\sum_{i=1}^8 w_i^k = 1
\end{cases}$$

By solving the above model, we calculate the PNS of linguistic terms with respect to attributes $r_j (j = 1, 3, 4, 6, 7, 8)$ for four decision makers $e_k (k = 1, 2, 3, 4)$(see Tables 10–15). Fig. 5 shows the obtained PNS graphically.

From Tables 10–15 and Fig. 5, the computed PNS values differ between the decision makers, which reflects the PIS of decision makers.

Table 16 lists the personalized attribute wights associated with decision makers $e_k (k = 1, 2, 3, 4)$, which are distinct for different decision makers.

6. Comparative study

In this section, we provide a comparison between the data-driven PIS learning model and the existing classical methods: the CWW based triangular membership function [14,29,6] and the 2-tuple linguistic model [14].

6.1. A comparison index for comparing the three methods

In the CWW model based on a triangular membership function and the 2-tuple linguistic model, the semantics for decision makers are represented by triangular fuzzy numbers and linguistic 2-tuples, respectively, and we assume that in these two methods, the decision makers have the same semantics regarding words. In the CWW approach based on a triangular membership function [14,29,6], the semantics of linguistic terms $s_i (i = 0, 1, ..., g)$ for decision makers are as follows:

$$A(s_i) = \begin{cases} (0, 0, \frac{1}{g}), i = 0 \\ (\frac{i-1}{g}, \frac{i}{g}, \frac{i+1}{g}), i = 1, 2, ..., g-1 \\ (\frac{i-1}{g}, 1, 1), i = g \end{cases}$$

In the CWW model based on a 2-tuple linguistic model [14], the semantics for linguistic terms s_i for decision makers are set as $\Delta^{-1}(s_i) = \frac{i}{g}$ for $i = 0, 1, ..., g$, that is $\{0, \frac{1}{g}, \frac{2}{g}, ..., 1\}$.

To carry out the comparison, our goal is to provide an index that can compare the three methods: the CWW model based on

Table 9
Summary of dataset of house information in Den Bosch.

Dataset	Decision makers	#Alternatives	#Attributes	#Classes	Distribution of classes
House evaluation in Den Bosch	e_1	30	8	2	14/16
	e_2	30	8	2	15/15
	e_3	29	8	2	14/15
	e_4	30	8	2	16/14

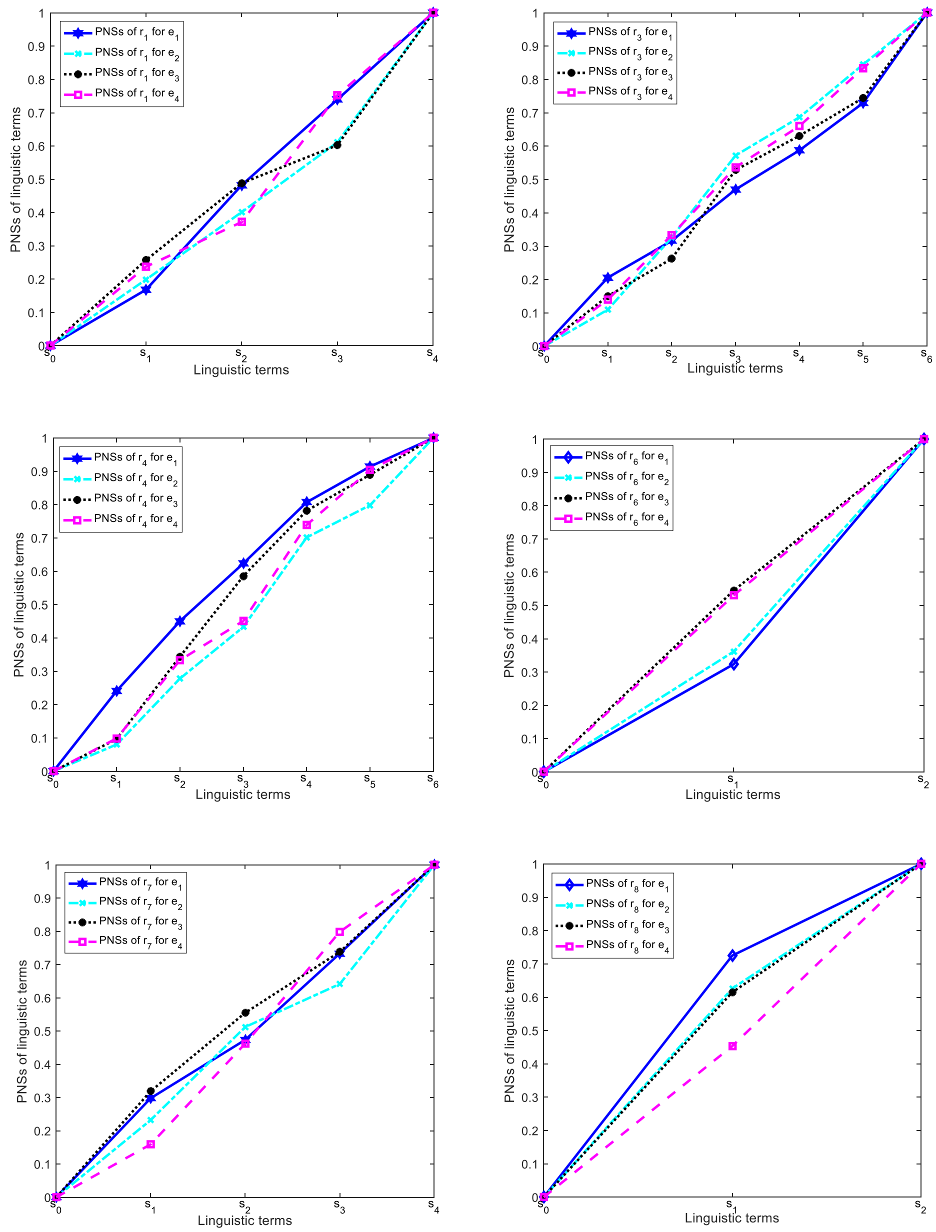

Fig. 5. PNS of linguistic terms with respect to $r_t(t = 1, 3, 4, 6, 7, 8)$ for $e_k(k = 1, 2, 3, 4)$.

Table 10
$PNS^k(s_t^1)$ with respect to attribute r_1 for $k = 1, 2, 3, 4$ and $t = 0, 1, 2, 3, 4$.

	$PNS^k(s_0^1)$	$PNS^k(s_1^1)$	$PNS^k(s_2^1)$	$PNS^k(s_3^1)$	$PNS^k(s_4^1)$
$k = 1$	0	0.1683	0.4822	0.7395	1
$k = 2$	0	0.1987	0.4012	0.611	1
$k = 3$	0	0.2571	0.4883	0.6024	1
$k = 4$	0	0.2373	0.3716	0.7523	1

a triangular membership function, the CWW model based on a 2-tuple linguistic model, and our proposed data-driven PIS learning model. To realize this idea, we propose a comparison index to measure the inconsistency/error between the alternatives' output values and their respective classes.

Let $C^k = \{C_1^k, ..., C_h^k\}$ with $C_i^k \succ C_{i+1}^k$ for $i = 1, 2, ..., h-1$ be a set of alternative classes. According to Eq. (7), let $F^k(x_i)$ be the comprehensive evaluation value of alternative x_i associated with deci-

Table 11
$PNS^k(s_t^3)$ with respect to attribute r_3 for $k = 1, 2, 3, 4$ and $t = 0, 1, 2, 3, 4, 5, 6$.

	$PNS^k(s_0^3)$	$PNS^k(s_1^3)$	$PNS^k(s_2^3)$	$PNS^k(s_3^3)$	$PNS^k(s_4^3)$	$PNS^k(s_5^3)$	$PNS^k(s_6^3)$
$k = 1$	0	0.2057	0.3163	0.4701	0.5873	0.7299	1
$k = 2$	0	0.1097	0.3263	0.572	0.6866	0.8452	1
$k = 3$	0	0.1503	0.2625	0.5285	0.6311	0.7447	1
$k = 4$	0	0.1398	0.3335	0.5367	0.6601	0.8339	1

Table 12
$PNS^k(s_t^4)$ with respect to attribute r_4 for $k = 1, 2, 3, 4$ and $t = 0, 1, 2, 3, 4, 5, 6$.

	$PNS^k(s_0^4)$	$PNS^k(s_1^4)$	$PNS^k(s_2^4)$	$PNS^k(s_3^4)$	$PNS^k(s_4^4)$	$PNS^k(s_5^4)$	$PNS^k(s_6^4)$
$k = 1$	0	0.2413	0.4505	0.6238	0.8073	0.9142	1
$k = 2$	0	0.0805	0.2783	0.4331	0.7017	0.7983	1
$k = 3$	0	0.097	0.3441	0.5856	0.7819	0.8898	1
$k = 4$	0	0.0988	0.3335	0.4513	0.7391	0.9044	1

Table 13
$PNS^k(s_t^6)$ with respect to attribute r_6 for $k = 1, 2, 3, 4$ and $t = 0, 1, 2$.

	$PNS^k(s_0^6)$	$PNS^k(s_1^6)$	$PNS^k(s_2^6)$
$k = 1$	0	0.3237	1
$k = 2$	0	0.3619	1
$k = 3$	0	0.5454	1
$k = 4$	0	0.5307	1

Table 14
$PNS^k(s_t^7)$ with respect to attribute r_7 for $k = 1, 2, 3, 4$ and $t = 0, 1, 2, 3, 4$.

	$PNS^k(s_0^7)$	$PNS^k(s_1^7)$	$PNS^k(s_2^7)$	$PNS^k(s_3^7)$	$PNS^k(s_4^2)$
$k = 1$	0	0.2983	0.4737	0.7332	1
$k = 2$	0	0.2326	0.5123	0.6416	1
$k = 3$	0	0.3193	0.5551	0.7383	1
$k = 4$	0	0.1595	0.4623	0.7984	1

Table 15
$PNS^k(s_t^8)$ with respect to attribute r_8 for $k = 1, 2, 3, 4$ and $t = 0, 1, 2$.

	$PNS^k(s_0^8)$	$PNS^k(s_1^8)$	$PNS^k(s_2^8)$
$k = 1$	0	0.7252	1
$k = 2$	0	0.6264	1
$k = 3$	0	0.6151	1
$k = 4$	0	0.4537	1

sion maker e_k, where

$$F^k(x_i) = \sum_{j=1}^{q} w_j^k \times f(l_{ij}^k) \tag{18}$$

with $f(l_{ij}^k)$ representing the semantics of the element l_{ij}^k. That is, in our proposed model and in the CWW model based on triangular membership function and the 2-tuple linguistic model, we have $f(l_{ij}^k)$ equal to $PNS^k(l_{ij}^k)$, $A(l_{ij}^k)$, and $\Delta^{-1}(l_{ij}^k)$, respectively.

For the classes $C_i^k \succ C_{i+1}^k$ in C^k, the comprehensive value of the alternatives in class C_i^k should be larger than that in class C_{i+1}^k, i.e., $F^k(x_a) \geq F^k(x_b)$ for $x_a \in C_i^k$ and $x_b \in C_{i+1}^k$, but if $F^k(x_a) < F^k(x_b)$, it indicates an inconsistency between the comprehensive values of alternatives x_a and x_b and their classes C_i^k and C_{i+1}^k.

To complete a comparison between triangular membership values, we use the method proposed in [30]: let the mean of $A(s_a)$ be $\mu(A(s_a)) = \frac{1}{3}(l(s_a) + m(s_a) + n(s_a))$ and the standard deviation$\sigma(A(s_a)) = \frac{1}{18}(l^2(s_a) + m^2(s_a) + n^2(s_a) - l(s_a)m(s_a) - l(s_a)n(s_a) - m(s_a)n(s_a))$, using the coefficient of variation (CV) index to compare triangular fuzzy numbers, we have $CV(A(s_a)) = \frac{\sigma(A(s_a))}{|\mu(A(s_a))|}$ with $\mu \neq 0$ and $\sigma > 0$ and the triangular fuzzy number with smaller value is ranked higher. Here, $A(s_a) > A(s_b)$, if $CV(A(s_a)) < CV(A(s_b))$.

From both F^k and C^k, we can obtain the comparison results between alternatives and estimate whether one alternative is superior or inferior to another, both being associated with the same decision maker e_k; thus, the error/inconsistency between the alternative performance reflected from F^k and C^k should be as small as possible. We use the following comparison index:

$$CI(F^k, C^k) = \frac{1}{\sum\limits_{i<j} \#C_i^k \cdot \#C_j^k} \sum_{i<j} \sum_{(x_q, x_l) \in C_i^k \times C_j^k} S(x_q, x_l) \tag{19}$$

$$where S(x_q, x_l) = \begin{cases} 1, \text{if} F^k(x_q) < F^k(x_l) \\ 0, otherwise \end{cases} \tag{20}$$

6.2. Comparative analysis

By solving Eq. (19) with the car evaluation dataset and house evaluation dataset described in Section 5, we obtain the value of $CI(F^k, C^k)$, and we consider this as the comparison index for the above three methods. The smaller the value of $CI(F^k, C^k)$, the better the performance of the model.

Tables 17 and 18 present the values of $CI(F^k, C^k)(k = 1, 2, 3, 4)$ for the three different methods based on the two datasets, respectively. Fig. 6 is a graphical representation of the values of $CI(F^k, C^k)(k = 1, 2, 3, 4)$.

The following observations can be drawn from Tables 17 and 18 and Fig. 6: by comparing the obtained values $CI(F^k, C^k)(k = 1, 2, 3, 4)$, we can see that the error/inconsistency between the comprehensive values of alternatives and their associated classes by using our proposed data-driven PIS learning model is much smaller than other two methods without implementing PIS, which

Table 16
Attribute weights W^k associated with e_k for $k = 1, 2, 3, 4$.

	w_1^k	w_2^k	w_3^k	w_4^k	w_5^k	w_6^k	w_7^k	w_8^k
$k = 1$	0.1075	0.2077	0.0839	0.2286	0.13	0.0589	0.0704	0.113
$k = 2$	0.1091	0.2202	0.0741	0.1896	0.0632	0.1387	0.0645	0.1405
$k = 3$	0.071	0.082	0.1921	0.2127	0.1079	0.1785	0.0663	0.0895
$k = 4$	0.0837	0.1574	0.129	0.1946	0.1083	0.1994	0.0621	0.0655

Table 17
Values of $CI(F^k, C^k)(k = 1, 2, 3, 4)$ based on car evaluation dataset.

	$CI(F^1, C^1)$	$CI(F^2, C^2)$	$CI(F^3, C^3)$	$CI(F^4, C^4)$
Method based on triangular fuzzy numbers	0.5195	0.4961	0.4363	0.4524
Method based on linguistic 2-tuples	0.1523	0.1952	0.1186	0.0849
Data-driven PIS learning model	0.0807	0.0854	0.0244	0.0454

Table 18
Values of $CI(F^k, C^k)(k = 1, 2, 3, 4)$ based on house evaluation dataset.

	$CI(F^1, C^1)$	$CI(F^2, C^2)$	$CI(F^3, C^3)$	$CI(F^4, C^4)$
Method based on triangular fuzzy numbers	0.0792	0.0472	0.0589	0.0686
Method based on linguistic 2-tuples	0.0095	0.0183	0.0137	0.0061
Data-driven PIS learning model	0.0073	0.0056	0.0048	0.0017

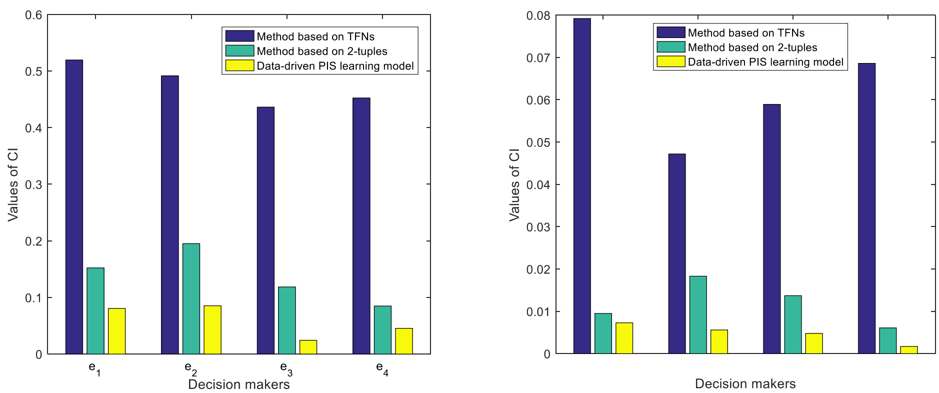

Fig. 6. Values of $CI(F^k, C^k)(k = 1, 2, 3, 4)$based on car evaluation dataset and house evaluation dataset, respectively.

indicates that the implementation of PIS can improve the compatibility of computed values of alternatives and their classes in LMADM.

7. Conclusions

In this study, considering the fact that words mean different things to different people, we propose a PIS learning method for DD-LMADM based on multi-attribute linguistic information and information on the classification of alternatives. The proposed model starts with a PIS multi-attribute learning function to provide a general computation form for evaluating the comprehensive values of alternatives in LMADM models; based on this function, we propose a data-driven PIS learning model that considers the distinct linguistic term sets and the unknown weights associated with the criteria to calculate the PNS of linguistic terms for decision makers.

The practical examples with car evaluation datasets and house evaluation datasets illustrate the applicability of the proposed data-driven PIS learning approach. Furthermore, we compare our approach with the existing CWW methods based on a triangular membership function and a 2-tuple linguistic model in terms of the inconsistency between the calculated preference values of alternatives and their class assignment.

Real world decision-making problems usually take place in complex environments [39,40,34], such as large-scale social networks, involving heterogenous linguistic preferences and flexible linguistic expressions. In future research, it would be an interesting direction to study data-driven PIS learning models in complex linguistic decision-making environments. In this study, we assume that the set of classification labels are same to each decision maker in DD-LMADM, and this situation is common in some practical decision problems (e.g., Tripadvisor.com). It will be an interesting future research to introduce personalized classification labels in DD-LMADM.

Author statement

The manuscript entitled "Data-driven method to learning personalized individual semantics to support linguistic multi-attribute decision making" is the authors' original work and has not been published nor has it been submitted simultaneously elsewhere. All authors have checked the manuscript and have agreed to its submission to the journal Omega.

Acknowledgments

This work was supported in part by the NSF of China under Grant 71901182, Grant 71601133 and Grant 71871149; in part by the Sichuan University under Grants YJ201906 and 2019-Business-C02; in part by the Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme; in part by the Southwest Jiaotong University under Grants YJSY-DSTD201918 and 2682021ZTPY073; in part by the China Postdoctoral Science Foundation under Grants 2020M673283 and 2021T140570.

References

[1] Durbach IN, Stewart TJ. Modeling uncertainty in multi-criteria decision analysis. Eur J Oper Res 2012;223(1):1–14.
[2] Lang MAK, Cleophas C, Ehmke JF. Multi-criteria decision making in dynamic slotting for attended home deliveries. Omega 2021;102:102305.
[3] Wallenius J, Dyer JS, Fishburn PC, Steuer RE, Zionts S, Deb K. Multiple criteria decision making, multiattribute utility theory: recent accomplishments and what lies ahead. Manag Sci 2008;54(7):1336–49.
[4] Wu S, Wu M, Dong Y, Liang H, Zhao S. The 2-rank additive model with axiomatic design in multiple attribute decision making. Eur J Oper Res 2020;287(2):536–45.
[5] Martínez L, Rodríguez RM, Herrera F. The 2-tuple linguistic model: computing with words in decision making. New York: Springer; 2015.
[6] Zadeh LA. The concept of a linguistic variable and its applications to approximate reasoning. *Inf Sci* Part I 1975;8:199–249 Part II, 8: 301-357; Part III, 9: 43-80.
[7] Zadeh LA. Fuzzy logic = computing with words. IEEE Trans Fuzzy Syst 1996;4:103–11.

[8] Mendel JM, Wu D. Perceptual computing: aiding people in making subjective judgments. New York, NY, USA: IEEE-Wiley; 2010.
[9] Mendel JM, Zadeh LA, Trillas E, Yager RR, Lawry J, Hagras H, Guadarrama S. What computing with words means to me: discussion forum. IEEE Comput Intell Mag 2010;5(1):20–6.
[10] Porro O, Pardo-Bosch F, Sánchez M, Agell N. Perceptual maps to aggregate information from decision makers. Artif Intell Res Dev 2021;339:37–45.
[11] Li CC, Dong YC, Herrera F, Herrera-Viedma E, Martínez L. Personalized individual semantics in computing with words for supporting linguistic group decision making. An application on consensus reaching. Inform Fusion 2017;33:29–40.
[12] Dong YC, Xu Y, Yu S. Computing the numerical scale of the linguistic term set for the 2-tuple fuzzy linguistic representation model. IEEE Trans Fuzzy Syst 2009;17(6):1366–78.
[13] Dong YC, Hong WC, Xu Y, Yu S. Numerical scales generated individually for analytic hierarchy process. Eur J Oper Res 2013;229(3):654–62.
[14] Herrera F, Martínez L. A 2-tuple fuzzy linguistic representation model for computing with words. IEEE Trans Fuzzy Syst 2000;8(6):746–52.
[15] Huang HY, Li CX. Extended personalized individual semantics with 2-tuple linguistic preference for supporting consensus decision making. IEICE Trans Inf Syst 2018;101:387–95.
[16] Li CC, Dong YC, Pedrycz W, Herrera F. Integrating continual personalized individual semantics learning in consensus reaching in linguistic group decision making. IEEE Trans Syst Man Cybern Syst 2022;52(3):1525–36.
[17] Tang X, Zhang Q, Peng Z, Yang S, Pedrycz W. Derivation of personalized numerical scales from distribution linguistic preference relations: an expected consistency-based goal programming approach. Neural Comput Appl 2019;31(12):8769–86.
[18] Tang X, Zhang Q, Peng Z, Pedrycz W, Yang S. Distribution linguistic preference relations with incomplete symbolic proportions for group decision making. Appl Soft Comput 2020;88:106005.
[19] Li CC, Gao Y, Dong Y. Managing ignorance elements and personalized individual semantics under incomplete linguistic distribution context in group decision making. Group Decis Negot 2021;30(1):97–118.
[20] Zhang HJ, Li C, Liu Y, Dong YC. Modelling personalized individual semantics and consensus in comparative linguistic expression preference relations with self-confidence: an optimization-based approach. IEEE Trans Fuzzy Syst 2021;29(3):627–40.
[21] Zhang HJ, Dong YC, Xiao J, Chiclana F, Herrera-Viedma E. Personalized individual semantics-based approach for linguistic failure modes and effects analysis with incomplete preference information. IISE Trans 2020;52(11):1275–96.
[22] Liang H, Li CC, Dong YC, Herrera F. Linguistic opinions dynamics based on personalized individual semantics. IEEE Trans Fuzzy Syst 2021;29(9):2453–66.
[23] Merigó JM, Palacios-Marqués D, Zeng S. Subjective and objective information in linguistic multi-criteria group decision making. Eur J Oper Res 2016;248(2):522–31.
[24] Díez, J., Luaces, Ó., Alonso-Betanzos, A., Troncoso, A., & Bahamonde Rionda, A. (2013). Peer assessment in MOOCs using preference learning via matrix factorization. In NIPS Workshop on Data Driven Education.
[25] Eric B, Freitas ND, Ghosh A. Active preference learning with discrete choice data. In: Proceedings of the advances in neural information processing systems; 2008. p. 409–16.
[26] Guo MZ, Zhang QP, Liao XW, Chen FY, Zeng DD. A hybrid machine learning framework for analyzing human decision-making through learning preferences. Omega 2021;101:102263.
[27] Liu JP, Kadziński M, Liao XW, Mao XX, Wang Y. A preference learning framework for multiple criteria sorting with diverse additive value models and valued assignment examples. Eur J Oper Res 2020;286(3):963–85.
[28] Liu JP, Kadziński M, Liao XW, Mao XX. Data-driven preference learning methods for value-driven multiple criteria sorting with interacting criteria. Inf J Comput 2021;33(2):586–606.
[29] Wang JH, Hao J. A new version of 2-tuple fuzzy linguistic representation model for computing with words. IEEE Trans Fuzzy Syst 2006;14(3):435–45.
[30] Cheng CH. A new approach for ranking fuzzy numbers by distance method. Fuzzy Sets Syst 1998;95(3):307–17.
[31] Dong YC, Li CC, Herrera F. Connecting the linguistic hierarchy and the numerical scale for the 2-tuple linguistic model and its use to deal with hesitant unbalanced linguistic information. Inf Sci 2016;367-368:259–78.
[32] Morente-Molinera JA, Perez IJ, Urena MR, Herrera-Viedma E. On multi-granular fuzzy linguistic modeling in group decision making problems: a systematic review and future trends. Knowl Based Syst 2015;74:49–60.
[33] Zhang Z, Guo C, Martínez L. Managing multigranular linguistic distribution assessments in large-scale multiattribute group decision making. IEEE Trans Syst Man Cybern Syst 2016;47(11):3063–76.
[34] Wu YZ, Dong YC, Qin J, Pedrycz W. Flexible linguistic expressions and consensus reaching with accurate constraints in group decision-making. IEEE Trans Cybern 2020;50(6):2488–501.
[35] Keeney RL, Raiffa H. Decision with multiple objectives: preferences and value tradeoffs. New York: John Wiley & Son; 1976.
[36] Chen YW, Wang CH, Lin SJ. A multi-objective geographic information system for route selection of nu clear waste transport. Omega 2008;36(3):363–72.
[37] Bohanec M, Rajkovic V. Expert system for decision making. Sistemica 1990;1(1):145–57.
[38] Daniels H, Kamp B. Application of MLP networks to bond rating and house pricing. Neural Comput Appl 1999;8(3):226–34.
[39] Gong ZW, Guo W, Herrera-Viedma E, Gong Z, Wei G. Consistency and consensus modeling of linear uncertain preference relations. Eur J Oper Res 2020;283(1):290–307.
[40] Wu ZB, Xu JP. Possibility distribution-based approach for MAGDM with hesitant fuzzy linguistic information. IEEE Trans Cybern 2016;46(3):694–705.