Information Sciences 321 (2015) 90–102

Contents lists available at ScienceDirect

Information Sciences

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

On learning of weights through preferences

Manish Aggarwal

Indian Institute of Management Udaipur, India

A R T I C L E   I N F O

Article history:
Received 21 September 2014
Received in revised form 8 May 2015
Accepted 22 May 2015
Available online 28 May 2015

Keywords:
Learning weights
Preference learning
Multi-criteria decision making
Learning-to-rank
Ordered weighted averaging
OWA

A B S T R A C T

We present a method to learn the criteria weights in multi-criteria decision making (MCDM) by applying emerging learning-to-rank machine learning techniques. Given the pairwise preferences by a decision maker (DM), we learn the weights that the DM attaches to the multiple criteria, characterizing each alternative. As the training information, our method requires the pairwise preferences of alternatives, as revealed by the DM. Once, the DM's decision model is learnt in terms of the criteria weights, it can be applied to predict his choices for any new set of alternatives. The empirical validation of the proposed approach is done on a collection of 12 standard datasets. The accuracy values are compared with those obtained for the state-of-the-art methods such as ranking-SVM and TOPSIS.

1. Introduction

Multi-criteria decision Making (MCDM) is a well-established subfield of operational research and decision sciences. It deals with the determination of the best choice among a set of alternatives. The decision making process in MCDM has the following three stages:

- evaluation of the various alternatives against multiple criteria,
- determination of the criteria weight vector that is specific to a DM,
- aggregation of the criteria evaluations taking into account the criteria weight vector.

An alternative $\boldsymbol{a}$ is preferred to $\boldsymbol{b}$ if the inequality relation $f(\boldsymbol{a}) > f(\boldsymbol{b})$ is true, where f is an aggregation function. The inequality $f(\boldsymbol{a}) > f(\boldsymbol{b})$ could be seen as a declaration of preference or desire of a DM, driven by the DM's subjective criteria evaluations of the various alternatives, and the relative importance values (weights) he attaches to various criteria. For example, an individual's preference of car $\boldsymbol{a}$ to car $\boldsymbol{b}$ is an outcome of a MCDM process in his cognition involving the criteria evaluations of the two cars, and the associated weight vector. Each DM may have a different set of criteria weights, as per his experience, priorities, and background. Our goal in this study is to infer the criteria weights from the given preferences of the form $\boldsymbol{a} \succ \boldsymbol{b}$, revealed by the DM, where $\boldsymbol{a}$ and $\boldsymbol{b}$ are the vectors of the criteria evaluations for alternatives $\boldsymbol{a}$ and $\boldsymbol{b}$, and $\succ$ indicates "preferred to".

E-mail address: magwal8@gmail.com

http://dx.doi.org/10.1016/j.ins.2015.05.034
0020-0255/

1.1. State-of-the-art and motivation

Many aggregation operators have appeared in the literature such as weighted arithmetic averaging [64], fuzzy weighted averaging [18,32], weighted geometric averaging [60], ordered weighted averaging (OWA) [63,68], induced ordered weighted averaging (IOWA) [69], generalized ordered weighted averaging (GOWA) [66], weighted ordered weighted averaging (WOWA) [53], and weighted induced OWA (WIOWA) [5]. A comprehensive overview of these operators can be found in [11,56,12,31,45].

More recently, a few aggregation operators have appeared specifically for practical decision making problems. For instance, in [43], aggregation operators are developed to aggregate linguistic values in MCDM. The class of compensative aggregation operators are proposed in [3,4] to consider the individual degrees of conjunctiveness or disjunctiveness. In [57,14], the prioritized weighted averaging operators are proposed for MCDM. A novel extension of the arithmetic mean, based on penalty functions that provides a representative output and satisfies idempotency, is proposed for non-convex intervals in [10]. In [47], the notion of a family of aggregation operators is conceived such that a set of aggregation operators are defined in the unit interval.

All these operators return an aggregation score for each alternative taking into account the weight vector of the DM. The identification of the weight vector corresponding to the multiple criteria is an important issue that has not drawn as much attention as it deserves. The earliest approaches for learning the weight vector are introduced in [72]. Some of the conventional approaches in this regard are Saaty's analytic hierarchy process (AHP) [48], SMART [20], centroid method [50], and a comparison based approach [49]. A parametric method is presented in [46]. A constrained optimization problem is solved to calculate the weight vector in [33]. In [34,22,28], the weight vector is obtained corresponding to the maximal entropy. Similarly, the authors in [29,7] derive a minimum variability weight vector. However, all these approaches are of limited applicability in practice as they need some additional information that is difficult to obtain in the real world.

This limitation is addressed in [63,65], where a method to deduce weights has been developed purely from the observed data, in terms of the criteria evaluations and the aggregation scores, using the Widrow–Hoff [59] approach. This approach is further extended in [23,24,54] to introduce a gradient descent technique for learning the weight vector, which has found applications in [69,55]. The central principle in these methods has been to minimize an error (difference between the actual and obtained aggregation scores) function, and hence they require an apriori knowledge of the actual aggregation scores, restricting their application in practice. For instance, it is difficult to expect the aggregation scores for different possible alternatives from a prospective car buyer.

Instead, it is far more convenient to have the relative preferences of the form $\boldsymbol{a} \succ \boldsymbol{b}$, on the basis of the multi-criteria evaluations, and such datasets are relatively easier to procure. In the present work, we are inspired to learn the weight vector on the basis of the observed preference information in the form $\boldsymbol{a} \succ \boldsymbol{b}$ as the training information. To this end, we apply the recent advances in preference learning [26] that is emerging as a new subfield of machine learning. Concomitantly, the proposed method helps to automate the learning process, and adds to the scalability.

1.2. The proposed work

Learning of the weight vector is an interesting problem, both from the MCDM and machine learning perspectives. While, MCDM is more focusing on the cognitive plausibility and interpretability of decision models, machine learning methods put more emphasis on the viewpoint of efficient algorithms and prediction performance. In the present work, we combine the MCDM models with the state-of-the-art machine learning methods for learning the weight vector, thereby cross-fertilizing these two complementary fields.

The two can significantly aid each other, as also shown in seminal papers in [17,19,30,16]. The present work is a concrete realization of this idea. Specifically, we construct the underlying preference model of a DM from his revealed preferences, by applying the recent algorithmic advances in preference learning (PL) methods. We provide the preference information in the form of pairwise comparisons between alternatives. Our learning algorithm is supposed to answer the following question: Given the alternative-criteria evaluations and the pairwise preferences revealed by a DM, what could be the criteria weight vector that this DM has in mind?

Once the DM's preference model is learnt in the form of the weight vector, it is then used to predict the DM's preferences for a new set of alternatives. The comparison of this prediction with the 'true ranking' offers a means for testing the performance of our method. Our methodology is very close to human cognition and reasoning. Humans including young children demonstrate the ability to make inferences about an individual's preferences on the basis of his choices [40]. The contributions are briefly summarized as:

- A PL-based learning model is introduced in Section 3 for learning criteria weights.
- Experimental setup and datasets are described in Section 4.
- Empirical validation on a set of 12 datasets is performed in Section 5.
- The prediction performance is compared with those from the state-of-the-art methods.