Neurocomputing 73 (2010) 2920–2927

Contents lists available at ScienceDirect

Neurocomputing

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

# A single-layer perceptron with PROMETHEE methods using novel preference indices

Yi-Chung Hu *

*Department of Business Administration, Chung Yuan Christian University, Chung-Li 32023, Taiwan, ROC*

ARTICLE INFO

*Article history:*
Received 19 September 2009
Accepted 4 August 2010
Communicated by G. Phillips-Wren

*Keywords:*
PROMETHEE methods
MCDA
Single-layer perceptron
Compromise operator
Preference relation

ABSTRACT

The Preference Ranking Organization METHods for Enrichment Evaluations (PROMETHEE) methods, based on the outranking relation theory, are used extensively in multi-criteria decision aid (MCDA). In particular, preference indices with weighted average aggregation representing the intensity of preference for one pattern over another pattern are measured by various preference functions. The higher the intensity, the stronger the preference is indicated. For MCDA, to obtain the ranking of alternatives, compromise operators such as the weighted average aggregation, or the disjunctive operators are often employed to aggregate the performance values of criteria. The compromise operators express the group utility or the majority rule, whereas the disjunctive operators take into account the strongly opponent or agreeable minorities. Since these two types of operators have their own unique features, it is interesting to develop a novel aggregator by integrating them into a single aggregator for a preference index. This study aims to develop a novel PROMETHEE-based single-layer perceptron (PROSLP) for pattern classification using the proposed preference index. The assignment of a class label to a pattern is dependent on its net preference index, which is obtained by the proposed perceptron. Computer simulations involving several real-world data sets reveal the classification performance of the proposed PROMETHEE-based SLP. The proposed perceptron with the novel preference index performs well compared to that with the original one.



## 1. Introduction

Traditional single-layer perceptrons (SLPs), trained by the backpropagation algorithm, with a single output neuron can find half planes bounded by a hyperplane. Researchers have applied SLPs to various two-class problems. Given two ordered classes $C_1$, $C_2$ such that $C_1 \succ C_2$, indicating that $C_1$ consists of the most preferred alternatives while $C_2$ consists of the least preferred alternatives. The sigmoid function commonly serves as the activation function of the output node and the desired output values of patterns in $C_1$ and $C_2$ can be specified as 1 and 0, respectively. A SLP is a utility function-based model which realizes criteria aggregation model based on absolute judgments. Then, the classification of the alternatives is based on the comparison of alternatives to a cut-off point. When the actual output value of an input pattern does not exceed a pre-specified cut-off point (e.g., 0.5), it can be assigned to $C_2$; otherwise, it can be assigned to $C_1$. Pattern with larger output value are more preferable for a dichotomous classification problems.

In addition to the utility function-based approach, outranking relations on the basis of pairwise comparisons are the other preference form in MCDA [8]. Given two patterns $\mathbf{a} \in C_1$ and $\mathbf{b} \in C_2$ for a multi-criteria classification problem from the viewpoint of the outranking relation theory (ORT) established by Roy [13], $C_1 \succ C_2$ indicates that **a** is at least as good as **b** [7,8,14]. The PROMETHEE methods introduced by Brans, Marechal, and Vincke [15–17,27] are extensively used and have played a significant role in multi-criteria analysis. In the PROMETHEE methods, the overall preference index $p(\mathbf{a}, \mathbf{b})$, which is specified as the weighted average aggregation of the preference of **a** over **b** in each criterion, is employed to measure the strength of the preference for **a** over **b** [14]. The outranking character over the other patterns and the outranked character by the other patterns for a pattern can be further estimated by the above preference indices. Previously, Doumpos and Zopounidis [14] developed a classification approach based on the above-mentioned outranking and outranked characters. They named this approach PAIRwise CLASsification (PAIRCLAS) and applied it to credit risk assessment. Actually, researchers have applied many members of the family of the PROMETHEE methods, especially the PROMETHEE I and II methods, to MCDA [9–12,18,19].

For MCDA, operators involving the combination of total and individual regret (or satisfaction) have been considered in the

* Tel.: +886 3 2655130.
*E-mail address:* ychu@cycu.edu.tw

0925-2312/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2010.08.002

overall aggregation on each criterion to rank alternatives from best to worst. The compromise operators (e.g., weighted average) can be employed to express the former related to group utility or majority rule, whereas the disjunctive operators are ones that can take account of the latter as related to strongly opponent or agreeable minorities. For instance, in order to compute the measure of closeness to the ideal solution for an alternative, the VIKOR method [28,29,41], developed from the $L_p$-metric in compromise programming [38,39], employs the weighted average ($L_1$) and the maximum operator ($L_\infty$) to provide majority rule and individual regret, respectively. The TOPSIS [26,32,39] could use a method different than VIKOR to balance total and individual satisfaction [29] by defining the closeness to the ideal solution for an alternative by determining the Euclidean distances ($L_2$) between this alternative and the ideal and negative ideal solutions. Following the ELECTRE methods introduced by Roy [6,13,33], Perny [7] presented the concordance and non-discordance principle for outranking methods to measure an overall verdict in favor of **a** and against **b**. It should be noted that Roy established the foundations of the ORT by developing the elimination and choice translating reality (ELECTRE) methods. When considering the overall preference index of PROMETHEE, since only majority rule is taken into account, it is an interesting prospect to develop a novel aggregator by incorporating total and individual satisfaction into a single aggregator for a preference index.

For SLPs, instead of using the traditional utility function-based transfer function, it could also be interesting to incorporate the above-mentioned novel preference index into the transfer function in view of the usefulness of the PROMETHEE methods. Thus, this study develops a novel SLP whose output neuron is represented by a set of connection weights and a PROMETHEE-based transfer function using the proposed preference index. The connection weights are interpreted as the relative importance of the respective criteria. By performing pairwise comparisons on an input pattern and all training patterns, the output value of the proposed SLP is the net preference index of the input pattern. To construct a PROMETHEE-based SLP (PROSLP) with high classification performance, this study employs genetic algorithms (GAs) [1–3] to develop a genetic-algorithm-based (GA-based) method that automatically determines the connection weights.

The rest of the paper is organized as follows. Sections 2 and 3 describe the concepts involved in the utility function-based models and PROMETHEE methods, respectively. Sections 4 and 5 demonstrate the framework of the proposed PROSLP and the GA-based learning algorithm, respectively. Section 6 reports the experimental results of the application of the proposed model to some real-world data sets. The results show that the PROSLP outperforms the traditional SLP and is comparable to the other fuzzy and non-fuzzy classification methods. Section 7 presents the discussion and conclusions.

## 2. Utility function-based models

Let $n$ be the number of criteria. Each pattern is a vector evaluated by $n$ attributes such that $\mathbf{x}_i=(x_{i1}, x_{i2},\ldots, x_{in})$ and $\mathbf{x}_j=(x_{j1}, x_{j2},\ldots, x_{jn})$. From the viewpoint of the multiple-attribute utility theory (MAUT), $U(\mathbf{x}_i)>U(\mathbf{x}_j)$ holds if and only if $\mathbf{x}_i$ is preferred to $\mathbf{x}_j$ (i.e., $\mathbf{x}_i \succ \mathbf{x}_j$), where $U(\mathbf{x}_i)$ and $U(\mathbf{x}_j)$ are the utilities of $\mathbf{x}_i$ and $\mathbf{x}_j$, respectively. Thus, the utility function describes the preference relation between patterns. Let $\mathbf{x}=(x_1, x_2,\ldots, x_n)$, the additive utility function is the most commonly used form:

$$U(\mathbf{x}) = \sum_{i=1}^{n} w_i x_i \tag{1}$$

where $0 \le w_1, w_2,\ldots, w_n \le 1$. These constants represent the relative weights of the attributes and sum up to one:

$$\sum_{i=1}^{n} w_i = 1 \tag{2}$$

Even although the sum of connection weights for traditional SLP may not be 1, the output value can also be interpreted as a synthetic evaluation or utility of the corresponding input pattern. As mentioned above, the larger the output value, the greater the possibility that an input pattern will be assigned to $C_1$. This means that the most preferred alternatives constitute $C_1$ and the least preferred alternatives constitute $C_2$. In other words, $C_1$ and $C_2$ are defined in an ordinal way (i.e., $C_1 \succ C_2$).

## 3. Preference index in PROMETHEE

ORT-based techniques provide preference information by performing pairwise comparisons between alternatives. Given $\mathbf{x}_i \in C_1$ and $\mathbf{x}_j \in C_2$, the ordering of the classes (i.e., $C_1 \succ C_2$) for the ORT implies that $\mathbf{x}_i$ is at least as good as $\mathbf{x}_j$. As mentioned above, the proposed method considers PROMETHEE methods involving the preference relation in the ORT [7]. The PROMETHEE family includes the methods of PROMETHEE I, II, III, IV, V and VI. The main differences between these methods are that PROMETHEE I gives a partial ranking of alternatives, version II allows a complete ranking with net flows, version III defines the preference and indifference relations by using the means and deviations of preference indices, version IV could deal with a set of infinite alternatives, version V is a procedure for multiple selection of alternatives under constraints, and version VI gives a representation of the human brain. Among the six methods, we select the PROMETHEE II method [16] since its concept of complete ranking is employed to develop the learning algorithm for the proposed SLP.

Although decision-makers can select many forms of criteria types (e.g., quasi criterion, level criterion, Gaussian criteria) for each criterion and certain preferential parameters must be specified for each function, for the sake of simplicity, this study uses Gaussian criterion as the preference function for each criterion. Gaussian criterion $H(d_k)$ ($1 \le k \le n$) ranging from 0 to 1 for $\mathbf{x}_i$ and $\mathbf{x}_j$ on criterion $k$ is defined as follows:

$$H(d_k) = 1 - e^{-d_k^2/2\sigma_k^2} \tag{3}$$

where $\sigma_k > 0$ is a preferential parameter that may be determined by decision-makers, and $d_k = x_{ik} - x_{jk}$. However, it may be difficult for decision-makers to specify a suitable value of $\sigma_k$. It should be noted that Olson [34] employed the data collected from major league professional baseball to examine the differences between the actual ranking and the ranking obtained by PROMETHEE. Olson's study found PROMETHEE II using Gaussian preference function was found to be particularly accurate.

Let $p_k$ be a one-dimensional valued preference relation such that a partial preference index $p_k(\mathbf{x}_i, \mathbf{x}_j) = H(d_k)$ for $x_{ik} \ge x_{jk}$, where $p_k(\mathbf{x}_i, \mathbf{x}_j) \in [0,1]$ indicates the intensity of the preference for $\mathbf{x}_i$ over $\mathbf{x}_j$ on criterion $k$. However, $p_k(\mathbf{x}_i, \mathbf{x}_j) = 0$ when $x_{ik} < x_{jk}$. Let the term $p(\mathbf{x}_i, \mathbf{x}_j) \in [0,1]$ denote an overall preference index which reveals the intensity of the preference for $\mathbf{x}_i$ over $\mathbf{x}_j$. $p(\mathbf{x}_i, \mathbf{x}_j)$ represents the flow from $\mathbf{x}_i$ to $\mathbf{x}_j$. The higher the $p(\mathbf{x}_i, \mathbf{x}_j)$ is, the stronger the preference for $\mathbf{x}_i$ over $\mathbf{x}_j$. $p(\mathbf{x}_i, \mathbf{x}_j)$ is derived by the weighted average of the intensity of the preference of $\mathbf{x}_i$ over $\mathbf{x}_j$ on each criterion:

$$p(\mathbf{x}_i,\mathbf{x}_j) = \sum_{k=1}^{n} w_k p_k(\mathbf{x}_i,\mathbf{x}_j) \tag{4}$$