Neurocomputing 179 (2016) 144–151

Contents lists available at ScienceDirect

Neurocomputing

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

Tolerance rough sets for pattern classification using multiple grey single-layer perceptrons

Yi-Chung Hu

Department of Business Administration, Chung Yuan Christian University, Chung Li Dist., Taoyuan City, Taiwan

ARTICLE INFO

Article history:
Received 7 June 2015
Received in revised form
13 November 2015
Accepted 25 November 2015
Communicated by: K Chan
Available online 17 December 2015

Keywords:
Tolerance rough set
Single-layer perceptron
Choquet integral
Grey relational analysis
Multiple attributes decision making (MCDM)

ABSTRACT

Tolerance rough sets (TRSs) can operate effectively on continuous attributes for pattern classification. The formulation of a similarity measure plays an important role for TRSs. The existence of certain relationships between any two patterns motivated us to use grey relational analysis (GRA) to implement a similarity measure on the basis of grey single-layer perceptrons (GSLPs). Additive and nonadditive GSLPs can perform additive and nonadditive versions of GRA, respectively. This paper contributes to use a one-class-in-one-network structure to construct the additive/nonadditive GSLP-based TRS for pattern classification by devoting each GSLP to one class. A GSLP-based tolerance class for each pattern can be generated by measuring the similarity for the output from the network. To yield a high classification performance of the proposed TRS-based classifier, a genetic-algorithm-based learning algorithm was designed to determine parameter specifications of the proposed classifier. Experimental results demonstrate that the test results of the proposed nonadditive classifier are better than, or comparable to, those of other known rough-set-based classification methods.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Rough set theory, which was introduced by Pawlak [13,14], is very useful for analyzing vague concepts [15–18]. For traditional rough-set-based methods, all quantitative attributes must be discrete [22]. However, discretization methods (e.g., [19,20]) can result in information losses, and there is no optimal discretization method for all decision problems [20]. Therefore, tolerance rough sets (TRSs) were developed to handle numerical attributes [21,22]. TRSs play an important role in pattern recognition [21–26].

In a traditional TRS, a tolerance relation is commonly defined by a simple distance measure [27] that indicates the proximity of any two patterns distributed in feature space. In [28], it was shown that the classification performance of a novel flow-based tolerance relation for pattern classification is superior to that of the traditional tolerance relation. In addition to the aforementioned measures for estimating proximity, since relationships exist between any two data sequences in the real world [11,29,30], it is interesting to construct a new similarity measure using relationships among patterns. In the field of MCDM, grey relational analysis (GRA) [29] is a technique appropriate for identifying relationships between a given reference sequence and several comparative sequences by viewing the reference sequence as the desired goal [30,31]. When each class is represented by a typical pattern, the grade of relationship between a common pattern (a comparative sequence) and such a typical pattern (a reference sequence) can easily be obtained through GRA [1]. Besides, the traditional GRA is additive, but an assumption of additivity may not be realistic in many applications [12] because attributes are not always independent of each other. Thus, the nonadditive version of a GRA should be taken into account when one considers an additive version. For pattern classification, this motivates us to incorporate GRA-like neural networks, which can consider the above-mentioned relationships among patterns and interactions among attributes, into the design of a new similarity measure for pattern classification.

On the basis of a single-layer perceptron (SLP), Hu [1] proposed a GRA-like neural network, named the additive/nonadditive grey single-layer perceptron (GSLP), for two-class problems. For GSLP, a corresponding typical pattern is specified for one class and a GRA is used to measure the degree of relationship between an input pattern, and thus a typical pattern. Additive and nonadditive GSLPs can perform additive and nonadditive versions, respectively, of a GRA for a given pattern. Note that the Choquet integral [2–5], which does not assume the independence of one element from another [7,8,12,44,45], is incorporated into the nonadditive GSLP.

The distinctive feature of this paper is to use a one-class-in-one-network structure to construct the new additive/nonadditive GSLP-based TRS (GSLP-TRS) for pattern classification by multiple GSLPs. Each GSLP in the network is devoted to one class. In the proposed TRS-based classifier, the similarity between any two

E-mail address: ychu@cycu.edu.tw

http://dx.doi.org/10.1016/j.neucom.2015.11.066
0925-2312/© 2015 Elsevier B.V. All rights reserved.

patterns can be measured by their respective outputs from the network. The GSLP-based tolerance classes are then constructed on the basis of the degree of proximity among patterns. Because genetic algorithms (GAs) are a powerful search and optimization method [9,10], we developed a GA-based method that automatically determines the relative weight of each attribute and a similarity threshold to achieve a high classification performance.

The rest of the paper is organized as follows. Section 2 briefly introduces TRS with a traditional similarity measure and a common classification procedure for a TRS-based classifier (TRSC). Section 3 describes additive/nonadditive GSLP. The proposed GSLP-based TRS is presented in Section 4. Section 5 describes a GA-based learning algorithm for constructing the proposed classifier using GSLP-based TRS. Section 6 reports experimental results for application of the proposed TRS-based classifier and other known rough-set-based classification methods to several real-world data sets. Sections 7 and 8 present the discussion and conclusions, respectively.

2. Tolerance rough sets

2.1. Traditional similarity measures

Let $\mathbf{x}_i\ R_a\ \mathbf{x}_j$ denote that $\mathbf{x}_i$ and $\mathbf{x}_j$ are similar with respect to attribute a, where R_a is a tolerance relation for attribute a. A standard similarity measure $S_a(\mathbf{x}_i, \mathbf{x}_j)$ with respect to R_a can be defined by a simple distance function as [27]:

$$S_a(\mathbf{x}_i,\mathbf{x}_j) = 1 - \frac{|\mathbf{x}_{ia} - \mathbf{x}_{ja}|}{\max_a - \min_a} \quad (1)$$

where $\mathbf{x}_{ia}$ and $\mathbf{x}_{ja}$ are the attribute values of $\mathbf{x}_i$ and $\mathbf{x}_j$ respectively, and $\max_a$ and $\min_a$ denote the maximum and minimum values respectively of the domain interval for attribute a. Of course, the same definition can be used for all attributes [22]. The relation between R_a and $S_a(\mathbf{x}_i, \mathbf{x}_j)$ is:

$$\mathbf{x}_i R_a \mathbf{x}_j \Leftrightarrow S_a(\mathbf{x}_i,\mathbf{x}_j) \ge \tau_a \quad (2)$$

where $\tau_a \in [0, 1]$ is the similarity threshold for attribute a. For A, an overall similarity measure $S_A(\mathbf{x}_i, \mathbf{x}_j)$ can be defined as:

$$S_A(\mathbf{x}_i,\mathbf{x}_j) = \frac{\sum_{a \in A} S_a(\mathbf{x}_i,\mathbf{x}_j)}{|A|} \quad (3)$$

The global tolerance relation R_A is related to $S_A(\mathbf{x}_i, \mathbf{x}_j)$ as follows:

$$\mathbf{x}_i R_A \mathbf{x}_j \Leftrightarrow S_A(\mathbf{x}_i,\mathbf{x}_j) \ge \tau \quad (4)$$

where $\tau \in [0, 1]$ is a global similarity threshold based on all attributes. A tolerance relation has reflexive and symmetric properties, but not transitivity property.

A tolerance class $TC(\mathbf{x}_i)$ of $\mathbf{x}_i$ can be generated for a certain τ by considering those patterns that have a tolerance relation with $\mathbf{x}_i$ as follows:

$$TC(\mathbf{x}_i) = \{\mathbf{x}_j \in U \mid \mathbf{x}_i R_A \mathbf{x}_j\} \quad (5)$$

X can be approximated by the lower approximation, $\underline{A_\tau}X$, and the upper approximation $\overline{A_\tau}X$. As in the traditional rough set, $\underline{A_\tau}X$ and $\overline{A_\tau}X$ can be defined by singletons as:

$$\underline{A_\tau}X = \{\mathbf{x} \mid \mathbf{x} \in U, TC(\mathbf{x}) \subseteq X\} \quad (6)$$

$$\overline{A_\tau}X = \{\mathbf{x} \mid \mathbf{x} \in U, TC(\mathbf{x}) \cap X \neq \phi\} \quad (7)$$

The tuple $\langle \underline{A_\tau}X, \overline{A_\tau}X \rangle$ is known as a TRS. In addition to singletons $\underline{A_\tau}X$ and $\overline{A_\tau}X$, it has been demonstrated that variants of approximations including subset and concept approximations can affect TRSC classification performance [28].

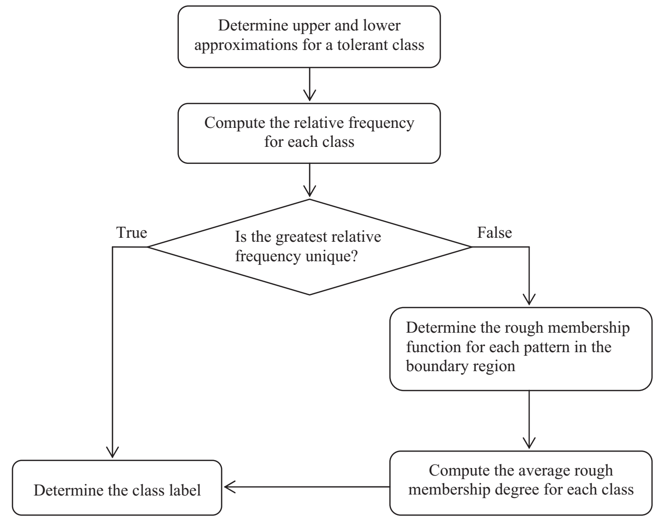

Fig. 1. A flow chart of TRSC for classifying a pattern.

2.2. TRS-based classifier

After tolerance classes have been determined for all patterns, a common classification procedure for a TRSC can be used to assign each pattern to a class. A flow chart of classifying a pattern $\mathbf{x}$ is illustrated in Fig. 1 based on [23,24]. Each main step is described as follows:

Step 1. Determine upper and lower approximations of a tolerant class

To acquire classification information with respect to $\mathbf{x}$, $\langle \underline{A_\tau}TC(\mathbf{x}), \overline{A_\tau}TC(\mathbf{x}) \rangle$ is determined at this step. $TC(\mathbf{x})$ is used because $\underline{A_\tau}TC(\mathbf{x})$ consists of patterns that are certainly similar to $\mathbf{x}$, whereas $\overline{A_\tau}TC(\mathbf{x})$ consists of patterns that are possibly similar to $\mathbf{x}$.

Step 2. Compute the relative frequency of each class by the lower approximation

If $\underline{A_\tau}TC(\mathbf{x})$ consists of at least two patterns, then the relative frequency of each decision class can be determined by $\underline{A_\tau}TC(\mathbf{x}) - \{\mathbf{x}\}$. The class label for $\mathbf{x}$ can be determined if the greatest relative frequency is unique; otherwise, it can be determined by the boundary region $\mathrm{BND}_A(TC(\mathbf{x}))$ of $\mathbf{x}$ (i.e., $\overline{A_\tau}TC(\mathbf{x}) - \underline{A_\tau}TC(\mathbf{x})$).

Step 3. Determine the rough membership function for each pattern in the boundary region

Because the patterns in $\underline{A_\tau}TC(\mathbf{x})$ have been considered in the previous step, only the patterns in $\mathrm{BND}_A(TC(\mathbf{x}))$ contribute to the classification in this step. Let X_l denote a set consisting of patterns belonging to the l-th class C_l $(1 \le l \le \alpha)$. For $\mathbf{y} \in \mathrm{BND}_A(TC(\mathbf{x})) \neq \phi$, the rough membership function $\mu_{C_l}(\mathbf{y})$ for the TRS with respect to A can be defined as follows:

$$\mu_{C_l}(\mathbf{y}) = \frac{|TC(\mathbf{y}) \cap X_l|}{|TC(\mathbf{y})|} \quad (8)$$

where $\mu_{C_l}(\mathbf{y}) \in [0, 1]$ and $|TC(\mathbf{y})|$ denotes the cardinality of $TC(\mathbf{y})$.

Step 4. Determine the average rough membership function of each class

Download English Version:

https://daneshyari.com/en/article/405970

Download Persian Version:

https://daneshyari.com/article/405970

Daneshyari.com