Neurocomputing 238 (2017) 56–66

Contents lists available at ScienceDirect

Neurocomputing

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

A modified fuzzy min–max neural network for data clustering and its application on pipeline internal inspection data

Jinhai Liu*, Yanjuan Ma, Huaguang Zhang, Hanguang Su, Geyang Xiao

College of Information Science and Engineering, Northeastern University, Shenyang 110819, China

ARTICLE INFO

Article history:
Received 18 August 2016
Revised 4 December 2016
Accepted 13 January 2017
Available online 27 January 2017

Communicated by H.R. Karimi

Keywords:
Clustering
Fuzzy min–max neural network
Hyperbox entropy
Internal inspection data

ABSTRACT

In this paper, a modified fuzzy min–max neural network (MFMC) for data clustering is proposed. In MFMC, the centroid information, the similarity and the noise of data are taken into the consideration. What's more, the hyperbox entropy (*HE*) is first introduced to evaluate the performance of each hyperbox when doing the contraction process. In addition, in order to test the performance of the MFMC model, a series of simulations on benchmark data sets are conducted. Then a real-world application study on the pipeline internal inspection data is also performed. The experimental result indicates that the MFMC has more excellent performance than other existed fuzzy min–max clustering algorithms.

1. Introduction

Unlike the pattern classification which needs labels to do the target classification, clustering is a method that needs little porior information to do the data analysis [1]. As an unsupervised method, clustering has an outstanding superiority of categorizing the unlabeled data into several groups or subsets [2]. The data samples in the same cluster have greater similarities than those in any other ones, for which it is widely used in many fields, such as data mining, pattern recognition, image segmentation, machine learning and so on [3].

Since the clustering has been intensively researched, it has attracted a lot of researchers' attention in recent years. Among so many clustering methods, fuzzy clustering and neural network clustering have been greatly developed.

In fuzzy clustering, an input pattern can belong to several clusters and the boundaries between the clusters are flexible [4], the theory has been broadly researched with fuzzy set [6] since it was first introduced into clustering by Zadeh in 1965 [5]. *The fuzzy sets* which is written by Zadeh is considered to be the fundamental of the fuzzy clustering. In addition, J. Bezdek's book *Pattern Recognition with Fuzzy Objective Functions* provides lots of new possibilities with clustering criteria, which plays an important role in the fuzzy clustering history [7]. It also has a great influence on the fuzzy model processing [8–10].

In addition, there are also some modeling approaches for neural network with Markov chain. They have been greatly researched by many researchers, such as Shen Hao, Wang Yueying and Wei,Y.L. [11–13]. As the neural network offers a biological support for clustering [14–24], the neural network provides an intrinsical structure for the clustering adaptively [25–30]. And it has a superiority in operating the parameters, such as the number, the size and the shape of clusters.

However, the fuzzy logic neural network has the advantages that both of the neural network and fuzzy logic have, which can deal with the uncertain information more capable than other learning algorithms [31]. Among fuzzy neural network methods, the fuzzy min–max neural network has been widely utilized since it was first introduced by Simpson [32,33]. The appealing attributes of the fuzzy min–max neural network are presented as follows:

(1) Little prior information is needed. Unlike fuzzy c-means (FCM) and k-means clustering algorithms which need to be pre-defined the number of clusters, there is no need to set the number of clusters in the fuzzy min–max neural network.
(2) The operation is simple. In the fuzzy min–max neural network, there are only two key parameters needing to be controlled [32,33]: one is the hyperbox size θ, the other is the sensitivity parameter γ. In addition, the MFMC is only an one-pass real-time online model. It does not need to divide into the training and testing process.

* Corresponding author.
E-mail addresses: jh_lau@126.com, liujinhai@ise.neu.edu.cn (J. Liu), myj_2016@163.com (Y. Ma), hgzhang@ieee.org (H. Zhang), suhanguang@sina.com (H. Su), xgyalan@outlook.com (G. Xiao).

http://dx.doi.org/10.1016/j.neucom.2017.01.036
0925-2312/

Fig. 1. The general procedure of the original fuzzy min–max learning algorithms.

(3) The boundaries are more flexible. Because of its fuzziness, the fuzzy neural network avoids the hard boundaries and makes the data analysis more easily and reasonably.

As the fuzzy neural network has so many appealing attributes, it has been applied in many fields. The mainly two fields are presented as follows:

(1) *Classification.* Since the fuzzy min–max neural network for pattern classification is proposed by Simpson in 1992, a lot of its variants have been proposed [34–39]. For example, the data-core-based fuzzy min–max neural network for pattern clustering(DCFMN), which not only deployed the membership function with two kinds of neurons, but also takes the noise into consideration [39].
(2) *Clustering.* As an excellent method, the fuzzy min–max neural network for data clustering(FMM) has been greatly developed: the general fuzzy min–max neural network for clustering and classification (GFMM) not only can do the classification, but also can do the clustering [34]. The stochastic fuzzy min–max neural network about reinforcement learning is proposed in 2001, which improves the performance of the original algorithm for clustering [40]. Moreover, a modified fuzzy min–max neural network for data clustering and its application on power quality monitoring (MFMM) is proposed in 2015 [41]. In MFMM, not only the previous hyperbox information is considered into the consideration, but also the centroid rule for hyperbox is created for hyperbox contraction process.

Motivated by the success of the FMM and its variants, a modified fuzzy min–max learning algorithm is proposed. It is called the modified fuzzy min–max neural network for data clustering and its application on pipeline internal inspection data (MFMC). The MFMC has following attributes:

(1) It is efficient in categorizing the data that has great similarities.
(2) An improved structure of the fuzzy min–max learning algorithm is designed, especially in the contraction process.
(3) The performance of each hyperbox is taken into the consideration. At the same time, an evaluating function called the *HE* is first proposed.
(4) Unnecessary overlap can be avoided after the overlap contraction.

In the rest of this paper, the review of the traditional fuzzy min–max learning algorithms is introduced in Section 2. Then the MFMC learning algorithm is introduced in Section 3, which contains the hyperbox selection, the overlap test and the contraction process. In Section 4, in order to evaluate the performance of the MFMC, many experimental comparisons are done with other clustering methods. Then a real-world application experiment on internal inspection data of the oil pipeline is conducted. At last, the conclusions are drawn in the last section.

2. The analysis of the fuzzy min–max clustering network

In this section, the original FMM and MFMM clustering algorithms are introduced. Then the analysis about them is presented in detail. The general procedure of the original fuzzy min–max learning algorithms is shown as below:

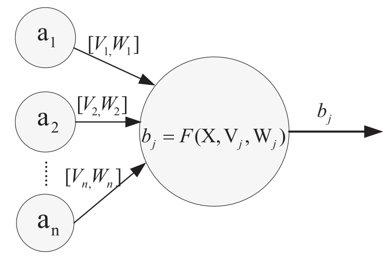

Fig. 2. The FMM membership function.

As is shown in Fig. 1, there are mainly three procedures in the whole fuzzy min–max clustering algorithms: the expansion process, the overlap test and the hyperbox contraction process. All of them will be introduced in greater detail in the following sections.

2.1. The FMM clustering network

The fuzzy min–max model is derived from the fuzzy sets [5]. Each hyperbox is regarded as a fuzzy set, which is determined by its minimum point V and maximum point W. The FMM neural network is presented in Fig. 2. The membership function B_j measures the degree that an input pattern belongs to a hyperbox, it is defined as follows:

$$B_j = (X_h, V_j, W_j, b(X_h, V_j, W_j)), \tag{1}$$

where $h = (1, 2, 3, \dots, n)$ is the serial number of the input pattern. And $X_h = (x_{h1}, x_{h2}, x_{h3}, \dots, x_{hn}) \epsilon I^n$ represents the hth input pattern. The $V_j = v_{j1}, v_{j2}, v_{j3}, \dots v_{jn}$ and $W_j = w_{j1}, w_{j2}, w_{j3}, \dots, w_{jn}$ are the minimum and maximum points of the jth hyperbox.

The membership value of jth hyperbox is b_j, which ranges from 0 to 1. If $b_j = 1$, it means that each dimension of the data sample X_h falls within the jth hyperbox. And the data sample is totally contained in the hyperbox. On the contrary, when the b_j is 0, it means that the data sample does not belong to the hyperbox at all. When the $b_j \epsilon [0, 1]$, the data sample that belongs to the hyperbox is in some extent. It shows that the nearer the input pattern approaches the hyperbox, the higher membership value it will obtain. b_j is described as follows:

$$b_j = (X_h, V_j, W_j) = \frac{1}{n}\sum_{i=1}[1 - f(X_{hi} - W_{ji}, \gamma) - f(V_{ji} - X_{hi}, \gamma)], \tag{2}$$

where $f(x, \gamma)$ is a two-parameter ramp threshold function.

$$f_{x,\gamma} = \begin{cases} 1 & if \quad x \times \gamma > 1 \\ x \times \gamma & if \quad 0 \le x \times \gamma \le 1 \\ 0 & if \quad x \times \gamma < 0 \end{cases} \tag{3}$$

As mentioned above, there are two key parameters in the FMM : one is the sensitivity parameter γ, which cotorls the descending speed of the membership function, the other is the expansion coefficient parameter θ.

From Fig. 2, it can be seen that there are three key steps in the FMM, they are expansion process, overlap test and contraction process. When there is an expansion process, the expansion criterion