Expert Systems with Applications 37 (2010) 1628–1634

Contents lists available at ScienceDirect

Expert Systems with Applications

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

A new computational intelligence technique based on human group formation

Arit Thammano [a,*], Jittraporn Moolwong [b]

[a] Computational Intelligence Laboratory, Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
[b] Kiatnakin Bank Plc., Sukhumvit 21, Sukhumvit Road, Bangkok 10110, Thailand

ARTICLE INFO

Keywords:
Computational intelligence
Human group formation
Classification
Data mining

ABSTRACT

This paper proposes a novel computational intelligence technique, based on the sociological concept of human group formation, with the aim to acquire a better solution to classification problems. The key concept of the human group formation is about the behavior of in-group members that try to unite with their own group as much as possible, and at the same time maintain social distance from the out-group members. This study compares the performance of the proposed model with that of fuzzy ARTMAP, radial basis function network, and learning vector quantization. Experimental results demonstrate the potential of the proposed approach in offering an efficient and effective solution to the problem.

1. Introduction

Computational intelligence is a fast-moving, multidisciplinary field. It covers various disciplines of computer science such as granular computing, neuro-computing, evolutionary computing, and artificial life (Konar, 2005). In recent years computational intelligence has attracted more and more attention over the traditional artificial intelligence. Unlike traditional artificial intelligence, computational intelligence is tolerant of imprecise information, partial truth, and uncertainty (Andina & Pham, 2007). Traditional artificial intelligence is very good in inductive and analogy-based learning (Konar, 2005) but it is inefficient to solve problems with large input sizes, as in data mining. The incompetence of traditional artificial intelligence has opened up new avenues for the non-conventional techniques like computational intelligence.

Classification, one of the most common data mining tasks, is the process of finding a set of models that describe and distinguish data classes or concepts, for the purpose of being able to use the model to predict the class of objects whose class label is unknown (Han & Kamber, 2001). A variety of computational intelligence techniques have been applied to deal with the classification problems such as artificial neural networks (Ham & Han, 1996; Huysmans, Baesens, Vanthienen, & Gestel, 2006; Lin, Xiao, & Micheli-Tzanakou, 1998; Loo & Rao, 2005; Mahanty & Dutta Gupta, 2004), fuzzy sets (Lee, Chen, Chen, & Jou, 2001; Li, Guo, & Kuo, 2005; Vernet & Kopp, 2002), and evolutionary algorithm (Chen & Hsu, 2006; Thammano & Trikeawcharoen, 2007; Wang & Wang, 2006; Zahiri & Seyedin, 2007). Among a number of computational intelligence methods in use, artificial neural networks are the most widely used approaches in solving classification problems. Many previous research works (Danaher et al., 1997; Goel et al., 2003; Lee et al., 1990; Quinlan, 1994; Russell & Norvig, 1995; Shavlik, Mooney, & Towell, 1991) show that neural network classifiers have better performance, lower classification error rate, and more robust to noise than other classification methods. Due to the superiority of neural networks in solving the classification problems, they were chosen to be reference methods in this research.

In this study, the new computational intelligence technique, which is inspired by the theory of human group formation, is proposed. The predictive performance of the proposed method is evaluated against three of the most powerful neural networks: the fuzzy ARTMAP neural network, the radial basis function network, and the learning vector quantization network.

This paper is divided into 6 sections. Following this introduction, Section 2 briefly introduces the three artificial neural networks which are used as a benchmark in this study. Section 3 presents the theoretical background of the human group formation. The proposed algorithm is described in Section 4. A brief description of the experimental data and the experimental results are given in Section 5. Finally, Section 6 is the conclusion.

2. Artificial neural networks

Artificial neural networks are the information processing systems that have been developed based on principles observed in human biological neural systems. Neural networks have many desirable characteristics such as resistance to noise, tolerance to distorted input patterns (ability to generalize), superior ability to recognize partially occluded or degraded images, ability to discriminate among overlapping pattern classes or classes with highly

* Corresponding author. Tel.: +66 2 723 4964; fax: +66 2 723 4910.
E-mail addresses: arit@it.kmitl.ac.th (A. Thammano), jittraporn@kiatnakin.co.th (J. Moolwong).

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.06.046

nonlinear boundaries, and potential for parallel processing (Pal & Pal, 2001).

Various types of artificial neural networks have made their appearance over the years. In this study, however, only three of the most powerful neural networks that are designed particularly for classification are chosen and briefly described.

2.1. Fuzzy ARTMAP

Fuzzy ARTMAP (Carpenter & Grossberg, 1992; Carpenter, Grossberg, Markuzon, Reynolds, & Rosen, 1992), introduced by Carpenter et al in 1992, is a supervised neural network with very interesting properties, such as a very fast convergence, a capacity of incremental learning of recognition categories, and multidimensional mapping in response to input vectors. However, its major drawback is that the network performance is affected by the ordering of training sample presentation (Carpenter et al., 1992; Loo & Rao, 2005). Carpenter, Grossberg, and Iizuka (1992) compared the performance of fuzzy ARTMAP with that of learning vector quantization (LVQ) and Backpropagation neural network (BPNN). The findings show that fuzzy ARTMAP has the edge over LVQ and BPNN in terms of prediction accuracy, and is far superior in terms of speed of learning.

As shown in Fig. 1, Fuzzy ARTMAP is a supervised neural network composed of two fuzzy ART modules (ART_a and ART_b). The two fuzzy ARTs are linked together via the map field F^{ab}. Each fuzzy ART module has two layers: F_1^a and F_1^b are the input layers of each module, while F_2^a and F_2^b are the category layers. During the training phase, the input vector and its desired output vector are presented to ART_a and ART_b, respectively. Before the input vector and its desired output vector are transmitted to their corresponding input layers F_1^a and F_1^b, in order to avoid the category proliferation problem, they are flowed through the complement coder where their strings are stretched to double the size by adding their complements. The complement coded vectors, $A = (a, a^c)$ and $B = (b, b^c)$, are then applied to the corresponding input layers. The ART_a and ART_b modules classify the input and desired output vector into categories, then the map field uses a vigilance criterion to evaluate whether ART_a category corresponds to ART_b category. The criterion is as follows:

$$\frac{\left|y^b \wedge w_J^{ab}\right|}{|y^b|} \geqslant \rho_{ab} \tag{1}$$

where y^b is the output vector of ART_b, J is the index of the winning node in F_2^a, w_J^{ab} is the weights of the connections from the Jth node in F_2^a to the map field F^{ab}, and ρ_{ab} is the vigilance parameter of the map field. If the criterion is not respected, the vigilance parameter of ART_a, ρ_a, will be increased by a minimum value just enough to force ART_a module to search for another category. However, when the vigilance criterion is respected, the weight vector is updated according to the equation:

$$w_J^{ab} = \beta_{ab} x^{ab} + (1 - \beta_{ab}) w_J^{ab} \tag{2}$$

where β_{ab} is the learning rate, which is between 0 and 1. x^{ab} is the output vector of F^{ab}.

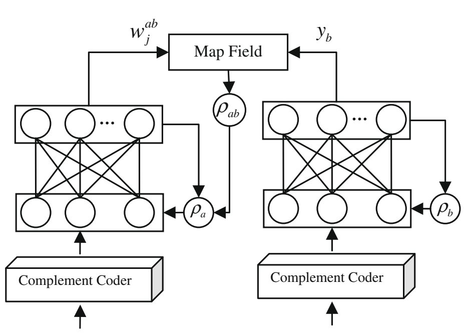

Fig. 1. The architecture of fuzzy ARTMAP.

2.2. Radial basis function network

Radial basis function network (RBF) (Fig. 2) comprises 3 layers: the input layer, the hidden layer, and the output layer. The input layer simply transfers the input patterns to the nodes in the hidden layer. Each node in the hidden layer then calculates its output as follows:

$$y_j = e^{\left(\frac{-\|x - v_j\|^2}{\sigma_j^2}\right)} \tag{3}$$

where $j = 1, 2, 3, \ldots, C$. It denotes the jth node in the hidden layer.

C is the total number of nodes in the hidden layer; x is the input vector; v_j is the center of the jth cluster; σ_j is the width of the jth cluster.

The outputs of the network are defined as the linear combination of the basis functions in the hidden layer.

$$y_k = \sum_{j=1}^{C} w_{jk} y_j \tag{4}$$

There are two steps in the training of the RBF network. First, the basis function parameters, v_j and σ_j, need to be determined. This is usually done by unsupervised learning algorithms such as K-means algorithm, fuzzy C-means algorithm, SOM, or the successive approximation method. Second, the connection weights between the hidden layer and the output layer, w_{jk}, are trained using a method like gradient descent method or least squares method.

2.3. Learning vector quantization network

Learning vector quantization network (LVQ) (Kohonen, 1990), a special case of the Kohonen self-organizing maps, is a supervised competitive neural network model in which each output node represents a particular class (Fig. 3). During training, the output node whose weight vector most closely matches the input pattern is chosen as the winner. If the winning node has the correct class label, its weight vector is moved toward the input pattern. However, if it belongs to the wrong class, its weight vector is moved away from the input.

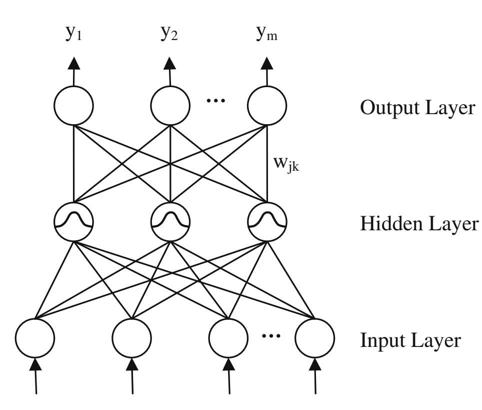

Fig. 2. The architecture of RBF network.