A New Evolutionary Neural Network Classifier

Arit Thammano* and Asavin Meengen**

Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang,
Bangkok, 10520 Thailand
arit@it.kmitl.ac.th*, asv_kmitl@hotmail.com**

Abstract. This paper proposes two new concepts: (1) the new evolutionary algorithm and (2) the new approach to deal with the classification problems by applying the concepts of the fuzzy c-means algorithm and the evolutionary algorithm to the artificial neural network. During training, the fuzzy c-means algorithm is initially used to form the clusters in the cluster layer; then the evolutionary algorithm is employed to optimize those clusters and their parameters. During testing, the class whose cluster node returns the maximum output value is the result of the prediction. This proposed model has been benchmarked against the standard backpropagation neural network, the fuzzy ARTMAP, C4.5, and CART. The results on six benchmark problems are very encouraging.

1 Introduction

In the past decades, data are being collected and accumulated at a dramatic pace. Therefore, there is an urgent need for a new generation of computational techniques and tools to assist humans in extracting useful information (knowledge) from the rapidly growing volumes of data [1]. This arouses many researchers to study into the area of data mining. One of the data mining functionalities, which plays an important role in business decision-making tasks, is classification. Classification 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 [2]. A variety of techniques have been applied to deal with the classification problems, such as neural networks, decision trees, and statistical methods. However, many previous research works show that neural network classifiers have a better performance, lower classification error rate, and more robust to noise than the other two methods mentioned above. The proposed evolutionary neural network classifier described in this paper employs the concept of the fuzzy c-mean clustering and the evolutionary algorithm to find and optimize the center and the standard deviation of each cluster. The performance of the proposed network is evaluated against the fuzzy ARTMAP, the backpropagation neural network, C4.5, and CART.

This paper is organized as follows. Following this introduction, section 2 presents the architecture of the evolutionary neural network classifier. The learning algorithm

T.B. Ho, D. Cheung, and H. Liu (Eds.): PAKDD 2005, LNAI 3518, pp. 249 – 255, 2005.

is described in section 3. In section 4, the experimental results are demonstrated and discussed. Finally, section 5 is the conclusions.

2 The Proposed Model

The architecture of the evolutionary neural network classifier is a three-layer feedforward neural network as shown in Figure 1. The first layer is the input layer, which consists of N nodes. Each node represents a feature component of the input data. The second layer is the cluster layer. The nodes in this second layer are constructed during the training phase; each node represents a cluster that belongs to one of the classes. The third layer is the output layer. Each node in the output layer represents a class. In this paper, the input vector is denoted by $X_i = (x_{i1}, \ldots, x_{iN})$, where i is the i^{th} input pattern, and N is the number of features in X. The nodes in the cluster layer are fully connected to the nodes in the input layer. Therefore, once the model receives the input and its associated target output (X_i, Y_i), the input vector X_i is directly transmitted to the cluster layer via these connections. Each node in the cluster layer then calculates the membership degree to which the input vector X_i belongs to its cluster j.

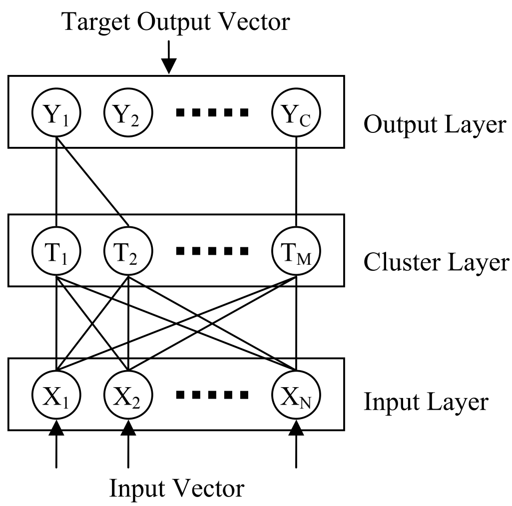

Fig. 1. Architecture of the evolutionary neural network classifier

$$T_j = \mu_j(X_i) = \text{Gaussian}(X_i, c_j, \sigma_j) = \exp\left[\frac{-d_{ji}^2}{2\sigma_j^2}\right]. \tag{1}$$

where j = 1, 2, ..., M. It denotes the j^{th} node in the cluster layer.