Available online at www.sciencedirect.com
ScienceDirect
Applied Soft Computing 8 (2008) 543–554
Applied Soft Computing
www.elsevier.com/locate/asoc

Fuzzy ARTMAP dynamic decay adjustment: An improved fuzzy ARTMAP model with a conflict resolving facility

Shing Chiang Tan [a,*], M.V.C. Rao [b], Chee Peng Lim [c]

[a] *Faculty of Information Science and Technology, Multimedia University, Malaysia*
[b] *Faculty of Engineering and Technology, Multimedia University, Malaysia*
[c] *School of Electrical and Electronic Engineering, University of Science Malaysia, Malaysia*

Received 5 December 2005; received in revised form 24 January 2007; accepted 22 March 2007
Available online 6 April 2007

Abstract

This paper presents a hybrid neural network classifier of fuzzy ARTMAP (FAM) and the dynamic decay adjustment (DDA) algorithm. The proposed FAMDDA model is a conflict-resolving classifier that can perform stable and incremental learning while settling overlapping of hyper-rectangular prototypes of different classes in minimizing misclassification rates. The performance of FAMDDA is evaluated using a number of benchmark data sets. The results are analyzed and compared with those from FAM and a number of machine learning classifiers. The outcomes show that FAMDDA has a better generalization capability than FAM, and its performance is comparable with those from other classifiers. The effectiveness of FAMDDA is also demonstrated in an application pertaining to condition monitoring of a circulating water system in a power generation station. Implications on the effectiveness of FAMDDA from the application point of view are discussed.

Keywords: Adaptive resonance theory; Classification; Dynamic decay adjustment; Fuzzy ARTMAP

1. Introduction

Data classification is one of the active domains of research which spans over a breadth of approaches and many years. It is a study concerned with separating distinct sets of observations into several groups and with assigning new observations to previously defined groups. Data classification is a prerequisite for many applications, such as pattern recognition, machinery fault detection and diagnosis, medical diagnosis, and business decision-making. Tremendous progress in this research area has resulted in a proliferation of different methods for building effective classifiers; to name a few, distance-based classifiers, such as the k-nearest neighbor classifier [1–4] that compute distance among the input vectors and then classify the input vectors into the individual class based on the smallest distance; kernel classifiers [5–7] that assign a class to an input pattern according to a majority vote among the labels of the training data using a kernel density estimator governed by a smoothing parameter; decision tree classifiers (e.g. ID3 [8] and C4.5 [9]) that form a structure of a tree for classification of an input data with an entropy-based measure; the mathematical models that are based on Bayesian decision theory [10,11]; artificial neural network (ANN)-based classifiers [12–18] that are trained using a given set of training data prior to predicting an output class of an unseen data sample.

In this paper, we focus on ANN-based classifiers. ANNs, inspired from brain modeling studies, are essentially models of biological neural systems that can perform computation on data. They have been successfully applied to many fields [19–22] with proven benefits in terms of efficiency and ability in solving complex problems. Many different ANN models have been developed to classify data in different applications, e.g., the radial basis function (RBF) network [12], self-organizing map (SOM) [13], higher order networks [14], adaptive resonance theory (ART) networks [15,16], probabilistic neural network (PNN) [17], and multilayer perceptron (MLP) with backpropagation [18]. Among the classifiers that offer a growing network structure include that family of ART neural networks. ART is an ANN model that reconciles the dilemma between stability and plasticity. ART models [15,16,23–28] can operate with an ability of recalling previously learned information (stability) without sacrificing an equally important ability of adapting to new information (plasticity).

* Corresponding author. Tel.: +60 6 2523195; fax: +60 6 2318840.
E-mail address: sctan@mmu.edu.my (S.C. Tan).

1568-4946/$ – see front matter
doi:10.1016/j.asoc.2007.03.006

The classification method examined in this paper is based on a supervised ART network, namely fuzzy ARTMAP (FAM) [26]. FAM processes uncertain (fuzzy) information and transforms it in terms of hyper-rectangles. Learning in FAM encompasses the recruitment of new hyper-rectangular prototypes and expansion of the boundary of existing prototypes in the feature space. Like other incremental ANNs, the growth criterion of FAM is subject to a similarity measure between the input pattern and the prototypes stored in the network. Given an input pattern, the prototype that has the highest degree of similarity with the input pattern is selected. A user-defined threshold is then used to decide whether or not the similarity level between the input pattern and the selected prototype is satisfactory to the user-defied level. If none of the prototypes can be found to meet the criterion, a new prototype is created and added to the network. However, a profound distinction between FAM and other incremental networks is that the fuzzy ART [25] modules of FAM undergo a two-stage hypothesis selection and test process. On presentation of an input pattern, a feedforward pass is carried out to identify the most similar prototype according to a competitive selection process. The winning prototype is then tested against a vigilance threshold in the feedback pass. As long as the vigilance criterion is not satisfied, a new cycle of search (selection and test) for a new winning prototype will be initiated. This search process is continued until the criterion is satisfied by an existing prototype, or the creation of a new node that includes the input pattern. Obviously, this feedback mechanism has attributed to the formation of stable yet plastic knowledge structure in FAM.

The growth process of FAM allows boundary expansion of existing prototypes as well as inclusion of new nodes to the network without retraining. One of the undesirable effects of prototypical growth in FAM is the overlapping effect extended from node expansion [26,31]. It may not be a problem when overlapping occurs among prototypes of the same class. However, if overlapping occurs among prototypes of different classes, it could bring an undesirable effect that causes the onset of ambiguity following hardly/undistinguishable regularities in the feature space. If the undesired overlapping prototypes of different classes remain unsettled, conflicting information would be retained as templates in the feature space. As a result, FAM might less likely to make accurate prediction/recognition based on overlapping prototypes. We believe that the existence of conflicting information in the network would incur negative impacts on its generalization performance. As such, a contraction procedure is necessary to reduce and/or resolve any undesired prototype overlaps of different classes so as to minimize misclassification. Since original FAM does not impart explicitly a learning scheme that is in settlement with a conflict, we therefore propose a novel conflict-resolving network based on FAM and the modified Dynamic Decay Adjustment (DDA) algorithm [29]. The resultant network is called FAMDDA. The main aims of the proposed method are to equip FAM with a conflict-resolving facility and to improve its generalization. Empirical experiments as well as real-world case studies are performed to assess the effectiveness of the proposed method.

The organization of this paper is as follows. In Section 2, related work on ANN classifiers which possess a width contraction facility is reviewed. In Section 3, details of the proposed hybrid FAM and DDA model are described. In Section 4, the effectiveness of the proposed FAMDDA is evaluated using several benchmark data sets. To demonstrate the applicability of FAMDDA to real-world applications, a condition monitoring task of a circulating water (CW) system in a power generation plant is investigated. Experiments as well as result analysis pertaining to the application of FAMDDA to monitoring the operating condition of the CW system are presented in Section 5. In Section 6, a summary concluding the work presented in this paper is given.

2. Conflict-resolving classifiers

The idea of deploying a contraction procedure in the network operation for eliminating overlaps is not a new one. One of the classifiers that undergoes prototypes expansion and contraction is the fuzzy min–max neural networks (FMM) [31]. A prototype in FMM is essentially an M-dimensional hyperbox fuzzy set in which each dimension is defined by a pair of minimum and maximum points and a sensitivity parameter that regulates the slope outside the two extreme points. The width of prototype is adjustable in the pattern space. To check whether or not an overlap exists between the hyperboxes of fuzzy sets of different classes, a dimension-by-dimension comparison among the hyperboxes is conducted according to several conditions as in Ref. [31]. If a conflict does occur, the smallest overlap along any dimension is eliminated. The conflict between the hyperboxes of fuzzy sets is then resolved. Note that FMM does not adjust each dimension of the overlapping hyperboxes. Instead, one dimension of the overlapping hyperboxes is adjusted in order to avoid a great change in the formation of the decision boundaries.

In retrospect, several conflict-resolving classifiers [31–33], developed to improve the generalization capability of the FMM model, were developed. The classifier proposed by Gabrys and Bargiela [32] (i.e., general fuzzy min–max (GFMM) neural network) is essentially an extension of FMM in which it combines supervised and unsupervised learning within a single learning scheme. Modifications have been made on the membership function, the adaptation of the hyperbox size and the expansion criterion of GFMM for coping with both labeled and unlabeled input patterns. Nonetheless, no significant modifications were made in the contraction procedure of GFMM. Hence, the width shrinking criteria are basically the same as in Ref. [31]. In Ref. [33], an ANFIS-like network (i.e., adaptive resolution min–max classifier (ARC)) and its pruning version (PARC) that are based on FMM are proposed. In their work, hyperboxes are created, and conflicts are resolved through a succession of cuts in addition to a fusion procedure that makes the coverage of the hyperboxes complete.

As for ART-based classifier, to our best knowledge, it is yet to have a contraction facility within the supervised ART framework. The proposed FAMDDA model inherits the advantages of its predecessors; it can perform stable and