Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 29 (2008) 392–406
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Improving multiclass pattern recognition with a co-evolutionary RBFNN

Minqiang Li *, Jin Tian, Fuzan Chen

School of Management, Tianjin University, Tianjin 300072, PR China

Received 30 March 2007; received in revised form 8 August 2007
Available online 5 November 2007

Communicated by W. Pedrycz

Abstract

A new hybrid scheme of the radial basis function neural network (RBFNN) model and the co-operative co-evolutionary algorithm (Co-CEA) is presented for multiclass classification tasks. This combination of the conventional RBFNN training algorithm and the proposed Co-CEA enforces the strength of both methods. First, the decaying radius selection clustering (DRSC) method is used to obtain the initial hidden nodes of the RBFNN model, which are further partitioned into modules of hidden nodes by the *K*-means method. Then, subpopulations are initialized on modules, and the Co-CEA evolves all subpopulations to find the optimal RBFNN structural parameters. Matrix-form mixed encoding and special crossover and mutation operators are designed. Finally, the proposed algorithm is tested on 14 real-world classification problems from the UCI machine learning repository, and experimental results illustrate that the algorithm is able to produce RBFNN models that have better prediction accuracies and simpler structures than conventional algorithms of classification.

Keywords: RBFNN; Co-operative co-evolutionary algorithms; *K*-means clustering; Multiclass classification

1. Introduction

Multiclass pattern recognition or multiclass classification tasks are ubiquitous in real-world, and there have been growing interests on the multiclass classification problems in the community of machine learning. A multiclass classification task attempts to learn a concept by the training instances with known labels in order to correctly label unknown instances.

There are basically two approaches to solve multiclass classification tasks. One is to create a group of pairwise classifications by matching one class with the others, and the total concept for the multiclass classification task consists of the pairwise classification concepts. The multiclass prediction is done based on the predictions of all two-class concepts. The other is to adapt directly learning algorithms to deal with multiclass problems. The first approach is feasible by restricting each instance taking only one label, and there are some well-known methods such as the error-correcting output codes (Dietterich and Bakiri, 1995), and round robin classification (Furnkranz, 2002). However, this approach fails to consider the correlations between the different labels of each instance in prediction (McCallum, 1999; Elisseeff and Weston, 2002). Nowadays, a lot of methods mainly belonging to the second approach have been proposed specially for multiclass learning tasks (Schapire and Singer, 2000; Kazawa et al., 2005).

The artificial neural network (ANN) has been verified to be a good method to learn multiclass classification concepts, and usually can yield the most predictive concepts for complicated problems. The mostly adopted network topology is radial basis function neural network (RBFNN)

* Corresponding author. Tel./fax: +86 22 27404796.
E-mail addresses: mqli@tju.edu.cn (M. Li), jtian_tju@yahoo.com.cn (J. Tian), fzchen@tju.edu.cn (F. Chen).

0167-8655/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2007.10.019

(Mitchell, 2003) due to a number of advantages compared with other types of ANNs, such as better prediction capabilities, simpler network structures, and faster learning process. Different variants of RBFNNs were invented to solve multiclass classification problems recently (Asim et al., 1995; Gao and Yang, 2002; Fu and Wang, 2003).

In this article, co-evolutionary RBFNN (CO-RBFNN) is proposed. It attempts to construct the RBFNN models for the multiclass classification problems by using a specially designed co-operative co-evolutionary algorithm (Co-CEA). The Co-CEA algorithm utilizes a divide-and-co-operative mechanism to evolve subpopulations with evolutionary algorithms in parallel (Zhao and Higuchi, 1996). After the initial hidden nodes are obtained by a decaying radius selection clustering (DRSC) method (Berthold and Diamond, 1995), a modified K-means method is employed to divide them further into modules. Then the hidden node modules are used to generate subpopulations for the Co-CEA to carry on the co-operative co-evolutionary searching. Collaborations among the modules are required to obtain complete solutions. The algorithm adopts a matrix-form mixed encoding which includes two determinant parameters of RBFNN's topology (the network centers and the radius widths) and a control vector. The optimal hidden layer structure is obtained by co-evolving all of the parameters. The CO-RBFNN is applied to 14 real-world classification problems from the UCI machine learning repository, and the CO-RBFNN achieves higher accuracies of prediction with a much simpler network structure in fewer evolutionary trials on nearly all of them.

The rest of the paper is structured as follows: Section 2 presents the fundamentals of our models such as the RBFNN architecture, the Co-CEA as well as the preliminaries of the multiclass learning. In Section 3, the proposed algorithm is described in details. Section 4 illustrates the new algorithm's performance on 14 UCI datasets in comparison with other learning techniques. Finally, Section 5 summarizes the key points of the paper and presents the concluding remarks.

2. Fundamentals

2.1. Multiclass learning

A typical classification problem of n classes based on N patterns or examples with known class membership is defined as follows. Let $S = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \ldots, (\boldsymbol{x}_N, y_N)\}$ be a set of N training samples, where $x_i \in \boldsymbol{R}^m$, and $y_i \in \boldsymbol{Y}$, $\boldsymbol{Y} = \{1, \ldots, n\}$. The multiclass learning is to output a multi-label classifier $h{:}\boldsymbol{X} \to 2^{|Y|}$ which optimizes some specific evaluation metric.

Researches on multiclass learning were initially motivated by the difficulty encountered in text categorization due to concept ambiguity, where each document may belong to several topics (labels) simultaneously (Zhang and Zhou, 2006). One famous approach to solving this problem is the Boostexter (Schapire and Singer, 2000), which is similar to the Adaboost (Freund and Schapire, 1997). Recently many approaches have been invented to solve the multiclass classification problems, such as the kernel-based mixture models (Xu et al., 2006), the decision tree (Freund and Mason, 1999), the Bayesian network (Kim and Ghahramani, 2006), and ANNs (Asim et al., 1995; Gao and Yang, 2002; Fu and Wang, 2003; Zhang and Zhou, 2006).

2.2. Co-operative co-evolutionary architecture

The co-evolutionary algorithm (CEA) is characterized with the adaptive fitness evaluation in co-evolutionary systems. There are two schemes of CEAs (Zhao and Higuchi, 1996): the competitive CEA and the co-operative CEA (Co-CEA). The co-operative scheme is adopted in this paper.

The Co-CEA consists of two or more interacting co-adapted subpopulations which are evolved independently by evolutionary algorithms. Each subpopulation contains individuals that represent a particular component of the problem solution, so that one member from each subpopulation is fetched to assemble a complete solution. In this mechanism, the fitness of an individual from a particular subpopulation is assessed by associating it with representatives from other subpopulations. There are many feasible ways for choosing representatives to do collaboration. For example, representatives from each subpopulation are chosen randomly, or they are selected based on their fitness. Generally, it is good to take the best individuals from subpopulations separately as the representatives that are stored in the elite pool.

The Co-CEA has been applied to problems such as function optimization (Iorio and Li, 2002), job-shop scheduling (Eriksson and Olsson, 1997) and concept learning (Potter and De Jong, 1998). Recently, it was used to train ANN, such as the co-operative co-evolutive approach for designing neural network ensembles (García-Pedrajas et al., 2005), and the Co-CEA architecture for supporting the dynamic creation of subpopulations in building cascade networks (Potter and De Jong, 2000).

In this paper, the RBFNN hidden structure is divided into multiple modules. Each module is used to initialize a subpopulation, and an individual, corresponding to a candidate of a module, represents a subcomponent of the total network structure. An individual in a subpopulation receives fitness based on how well it performs in conjunction with representatives from other subpopulations. Subpopulations are evolved by special designed evolutionary algorithms (EAs), and best individuals are output to form the complete solutions.

2.3. RBFNN

The RBFNN is a three-layer feed-forward neural network with multi-inputs, multi-outputs, and linear output mapping between the hidden nodes and the output nodes