Applied Soft Computing 13 (2013) 654–666

Contents lists available at SciVerse ScienceDirect

Applied Soft Computing

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

Meta-cognitive RBF Network and its Projection Based Learning algorithm for classification problems

G. Sateesh Babu, S. Suresh*

School of Computer Engineering, Nanyang Technological University, Singapore

ARTICLE INFO

Article history:
Received 2 February 2012
Received in revised form 24 May 2012
Accepted 31 August 2012
Available online 23 September 2012

Keywords:
Meta-cognitive learning
Self-regulatory thresholds
Radial basis function network
Multi-category classification
Projection Based Learning

ABSTRACT

'Meta-cognitive Radial Basis Function Network' (McRBFN) and its 'Projection Based Learning' (PBL) algorithm for classification problems in sequential framework is proposed in this paper and is referred to as PBL-McRBFN. McRBFN is inspired by human meta-cognitive learning principles. McRBFN has two components, namely the cognitive component and the meta-cognitive component. The cognitive component is a single hidden layer radial basis function network with evolving architecture. In the cognitive component, the PBL algorithm computes the optimal output weights with least computational effort by finding analytical minima of the nonlinear energy function. The meta-cognitive component controls the learning process in the cognitive component by choosing the best learning strategy for the current sample and adapts the learning strategies by implementing self-regulation. In addition, sample overlapping conditions are considered for proper initialization of new hidden neurons, thus minimizes the misclassification. The interaction of cognitive component and meta-cognitive component address the *what-to-learn*, *when-to-learn* and *how-to-learn* human learning principles efficiently. The performance of the PBL-McRBFN is evaluated using a set of benchmark classification problems from UCI machine learning repository and two practical problems, viz., the acoustic emission signal classification and the mammogram for cancer classification. The statistical performance evaluation on these problems has proven the superior performance of PBL-McRBFN classifier over results reported in the literature.

1. Introduction

Neural networks are powerful tools that can be used to approximate the complex nonlinear input–output relationships efficiently. Hence, from the last few decades neural networks are extensively employed to solve real world classification problems [1]. In a classification problem, the objective is to learn the decision surface that accurately maps an input feature space to an output space of class labels. Several learning algorithms for different neural network architectures have been used in various problems in science, business, industry and medicine, including the handwritten character recognition [2], speech recognition [3], biomedical medical diagnosis [4], prediction of bankruptcy [5], text categorization [6] and information retrieval [7]. Among various architectures reported in the literature, Radial Basis Function (RBF) network gaining attention due to its localization property of Gaussian function, and widely used in classification problems. Significant contributions to RBF learning algorithms for classification problems are broadly classified into two categories: (a) *Batch learning algorithms*: Gradient descent based learning was used to determine the network parameters [8]. Here, the complete training data are presented multiple times, until the training error is minimum. Alternatively, one can implement random input parameter selection with least square solution for the output weight [9,10]. In both cases, the number of Gaussian functions required to approximate the true function is determined heuristically. (b) *Sequential learning algorithms*: The number of Gaussian neurons required to approximate the input–output relationship is determined automatically [11–15]. Here, the training samples are presented one-by-one and discarded after learning. Resource Allocation Network (RAN) [11] was the first sequential learning algorithm introduced in the literature. RAN evolves the network architecture required to approximate the true function using novelty based neuron growth criterion. Minimal Resource Allocation Network (MRAN) [12] uses a similar approach, but it incorporates error based neuron growing/pruning criterion. Hence, MRAN determines compact network architecture than RAN algorithm. Growing and Pruning Radial Basis Function Network [13] selects growing/pruning criteria of the network based on the significance of a neuron. A sequential learning algorithm using recursive least squares presented in [14], referred as an On-line Sequential Extreme Learning Machine (OS-ELM). OS-ELM chooses input weights randomly with fixed number of hidden neurons and analytically determines the output weights using minimum norm least-squares. In case of sparse and imbalance data sets, the random

* Corresponding author. Tel.: +65 6790 6185.
E-mail address: ssundaram@ntu.edu.sg (S. Suresh).

1568-4946/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.asoc.2012.08.047

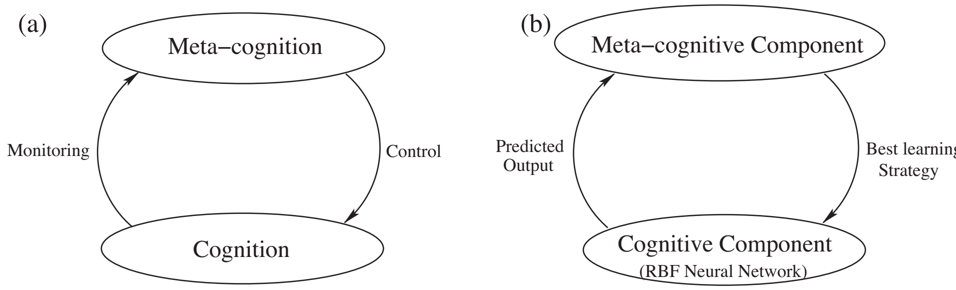

Fig. 1. (a) Nelson and Narens Model of meta-cognition and (b) McRBFN Model.

selection of input weights with fixed number of hidden neurons in the OS-ELM affects the performance significantly as shown in [16]. In neural-fuzzy framework, Evolving Fuzzy Neural Networks (EFuNNs) [17] is the novel sequential learning algorithm. It has been shown in [15] that the aforementioned algorithms works well for the function approximation problems than the classification problems. A Sequential Multi-Category Radial Basis Function network (SMC-RBF) [15] considers the similarity measure within class, misclassification rate and prediction error are used in neuron growing and parameter update criterion. SMC-RBF has been shown that updating the nearest neuron parameters in the same class as that of the current sample helps in improving the performance than updating a nearest neuron in any class.

Aforementioned neural network algorithms use all the samples in the training data set to gain knowledge about the information contained in the samples. In other words, they possess information-processing abilities of humans, including perception, learning, remembering, judging, and problem-solving, and these abilities are cognitive in nature. However, recent studies on human learning has revealed that the learning process is effective when the learners adopt self-regulation in learning process using meta-cognition [18,19]. Meta-cognition means *cognition about cognition*. In a meta-cognitive framework, human-beings think about their cognitive processes, develop new strategies to improve their cognitive skills and evaluate the information contained in their memory. If a radial basis function network analyzes its cognitive process and chooses suitable learning strategies adaptively to improve its cognitive process then it is referred to as 'Meta-Cognitive Radial Basis Function Network' (McRBFN). Such a McRBFN must be capable of deciding *what-to-learn*, *when-to-learn* and *how-to-learn* the decision function from the stream of training data by emulating the human self-regulated learning.

Self-adaptive Resource Allocation Network (SRAN) [20] and Complex-valued Self-regulating Resource Allocation Network (CSRAN) [21] address the *what-to-learn* component of meta-cognition by selecting significant samples using misclassification error and hinge loss error. It has been shown that the selecting appropriate samples for learning and removing repetitive samples helps in improving the generalization performance. Therefore, it is evident that emulating the three components of human learning with suitable learning strategies would improve the generalization ability of a neural network. The drawbacks in these algorithms are: (a) the samples for training are selected based on simple error criterion which is not sufficient to address the significance of samples; (b) the new hidden neuron center is allocated independently which may overlap with already existed neuron centers leading to misclassification; (c) knowledge gained from past samples is not used; and (d) uses computationally intensive extended Kalman filter for parameter update. Meta-cognitive Neural Network (McNN) [22] and Meta-cognitive Neuro-Fuzzy Inference System (McFIS) [23] address the first two issues efficiently by using three components of meta-cognition. However, McNN and McFIS use computationally intensive parameter update and does not utilize the past knowledge stored in the network. Similar works using meta-cognition in complex domain are reported in [24,25]. Recently proposed Projection Based Learning in meta-cognitive radial basis function network [26] addresses the above issues in batch mode except proper utilization of the past knowledge stored in the network and applied to solve biomedical problems in [27–29]. In this paper, we propose a meta-cognitive radial basis function network and its fast and efficient projection based sequential learning algorithm.

There are several meta-cognition models available in human physiology and a brief survey of various meta-cognition models are reported in [30]. Among the various models, the model proposed by Nelson and Narens in [31] is simple and clearly highlights the various actions in human meta-cognition as shown in Fig. 1(a). The model is analogous to the meta-cognition in human-beings and has two components, the cognitive component and the meta-cognitive component. The information flow from the cognitive component to meta-cognitive component is considered monitoring, while the information flow in the reverse direction is considered control. The information flowing from the meta-cognitive component to the cognitive component either changes the state of the cognitive component or changes the cognitive component itself. Monitoring informs the meta-cognitive component about the state of cognitive component, thus continuously updating the meta-cognitive component's model of cognitive component, including, 'no change in state'.

McRBFN is developed based on the Nelson and Narens meta-cognition model [31] as shown in Fig. 1(b). Analogous to the Nelson and Narens meta-cognition model [31], McRBFN has two components namely the cognitive component and the meta-cognitive component as shown in Fig. 1(b). The cognitive component is a single hidden layer radial basis function network with evolving architecture. The cognitive component learns from the training data by adding new hidden neurons and updating the output weights of hidden neurons to approximate the true function. The input weights of hidden neurons (center and width) are determined based on the training data and output weights of hidden neurons are estimated using the projection based sequential learning algorithm. When a neuron is added to the cognitive component, the input/hidden layer parameters are fixed based on the input of the sample and the output weights are estimated by minimizing an energy function given by the hinge loss error as in [32]. The problem of finding optimal weights is first formulated as a linear programming problem using the principles of minimization and real calculus [33,34]. The Projection Based Learning (PBL) algorithm then converts the linear programming problem into a system of linear equations and provides a solution for the optimal weights, corresponding to the minimum energy point of the energy function. The meta-cognitive component of McRBFN contains a dynamic model of the cognitive component, knowledge measures and self-regulated thresholds. Meta-cognitive component controls the learning process of the cognitive component by choosing one of the four strategies for each sample in the training data set. When a