Improved Evolutionary Extreme Learning Machines Based on Particle Swarm Optimization and Clustering Approaches

Luciano D. S. Pacifico, Informatics Center, Federal University of Pernambuco, Recife, Brazil
Teresa B. Ludermir, Informatics Center, Federal University of Pernambuco, Recife, Brazil

ABSTRACT

Extreme Learning Machine (ELM) is a new learning method for single-hidden layer feedforward neural network (SLFN) training. ELM approach increases the learning speed by means of randomly generating input weights and biases for hidden nodes rather than tuning network parameters, making this approach much faster than traditional gradient-based ones. However, ELM random generation may lead to non-optimal performance. Particle Swarm Optimization (PSO) technique was introduced as a stochastic search through an n-dimensional problem space aiming the minimization (or the maximization) of the objective function of the problem. In this paper, two new hybrid approaches are proposed based on PSO to select input weights and hidden biases for ELM. Experimental results show that the proposed methods are able to achieve better generalization performance than traditional ELM in real benchmark datasets.

Keywords: Artificial Neural Networks, Extreme Learning Machine, Hybrid Systems, Particle Swarm Optimization, Population Stereotyping, Selection Operator

1. INTRODUCTION

Artificial neural networks (ANNs) are known as universal approximators and computational models with remarkable properties such as adaptability, capacity of learning by examples and the ability to generalize data (Haykin, 1998).

Neural networks are of great use in pattern classification applications, and through a supervised learning perspective, they are considered a general method for constructing mappings between a group of sample vectors (training set) and the corresponding classes, allowing the classification of unseen data as one of the classes learned in the training process.

One of the most used ANN models is the well-known Multi-Layer Perceptron (MLP). The training process of MLPs for pattern classification consists of two main tasks: selection of an appropriate architecture for the problem and the adjustment of the connection weights of the network.

DOI: 10.4018/jncr.2012070101
Gradient-based learning strategies, such as Backpropagation (BP) and its variant Levenberg-Marquardt (LM-BP), have been extensively used in the training of MLPs, but these approaches are usually slower than required in learning, and may also get stuck in local minima (Zhu et al., 2005).

Extreme learning machine (ELM) was proposed as an efficient learning algorithm for single-hidden layer feedforward neural network (SLFN) (Huang et al., 2006). ELM increases the learning speed by means of randomly generating weights and biases for hidden nodes, differently from gradient-based approaches, which commonly tune iteratively the network parameters.

Although ELM is fast and presents good generalization performance, as the output weights are computed based on the prefixed input weights and hidden biases using the Moore-Penrose (MP) generalized inverse, there may exist a set of non-optimal input weights and hidden biases, and it might suffer from the overfitting as the learning model will approximate all training samples well.

Global search techniques, such as Tabu Search (TS) (Glover, 1986), Evolutionary Algorithms (EAs, like Genetic Algorithm - GA) (Eiben & Smith, 2003), Differential Evolution (DE) (Storn & Price, 1995; Storn & Price, 1997), Particle Swarm Optimization (PSO) (Kennedy & Eberhart, 1995; Kennedy & Eberhart, 2001; van den Bergh, 2002) and Group Search Optimization (GSO) (He et al., 2006; He et al., 2009), are widely used in scientific and engineering problems, and these strategies have been combined with ANNs to perform various tasks, such as connection weight initialization, connection weight training and architecture design.

In this paper, we present two new hybrid evolutionary approaches based on Particle Swarm Optimization technique to select input weights and hidden biases for Extreme Learning Machine neural network: PSO-ELM-CS$_2$ and GCPSO-CS$_2$. These methods are extensions from the PSO-ELM-CS$_1$ and GCPSO-ELM-CS$_1$ approaches, respectively, presented in Pacifico and Ludermir (2012). The Particle Swarm Optimization (PSO) consists of a stochastic global search originated from the attempt to graphically simulate the social behavior of a flock of birds looking for resources.

For the proposed methods, individuals in the PSO population were divided into groups by a clustering algorithm, following the idea of “population stereotyping” presented in Kennedy (2000). The lbest topology was adopted in a way that PSO particles will update according to individuals from its neighborhood. A selection operator was also applied to all strategies, based on the ideas of Angeline (1999).

Some evolutionary strategies have been adopted for the ELM context. Zhu et al. (2006) introduces a hybrid form of differential evolutionary (DE) algorithm to search for optimal input weights and hidden biases for ELM, called E-ELM to train SLFN with more compact networks.

Xu and Shu (2006) presented a new evolutionary ELM based on PSO for prediction task. In Saraswathi et al. (2011) a combination of Integer Coded Genetic Algorithm (ICGA) and Particle Swarm Optimization (PSO), coupled with the ELM has been used for gene selection and cancer classification, where the ICGA and PSO-ELM selected an optimal set of genes which are then used to build a classifier to develop an algorithm (ICGA_PSO_ELM) that could handle sparse data and sample imbalance.

In Cho and Lee (2007), an optimization method based on Bacterial Foraging (BF) algorithm was proposed to adjust the input weights and hidden biases for the ELM.

Lahoz et al. (2011) presented a bi-objective micro genetic ELM ($\mu$G-ELM) to generate the hidden weights and biases for ELM, and it also used a regression based strategy to select the appropriate number of hidden nodes.

In Silva et al. (2011a), the ELM was combined with Group Search Optimization (GSO) algorithm (GSO-ELM), and four different forms of handling individuals (members) that fly out of the search space bounds were used. The GSO was used to optimize the input weights and hidden biases for ELM, and also has found a more compact architecture than ELM for four of the six tested datasets.
Recent Methodology in Connectionist Systems
Advancing Artificial Intelligence through Biological Process Applications (pp. 94-106).
www.igi-global.com/chapter/recent-methodology-connectionist-systems/4974?camid=4v1a

Genetic-Algorithm-Based Optimization of Fragile Watermarking in Discrete Hartley Transform Domain
Sudipta Kr Ghosal and Jyotsna Kumar Mandal (2016). Handbook of Research on Natural Computing for Optimization Problems (pp. 103-127).