Chapter 34
Two Diverse Swarm Intelligence Techniques for Supervised Learning

Tad Gonsalves
Sophia University, Japan

ABSTRACT

Particle Swarm Optimization (PSO) and Enhanced Fireworks Algorithm (EFWA) are two diverse optimization techniques of the Swarm Intelligence paradigm. The inspiration of the former comes from animate swarms like those of birds and fish efficiently hunting for prey, while that of the latter comes from inanimate swarms like those of fireworks illuminating the night sky. This novel study, aimed at extending the application of these two Swarm Intelligence techniques to supervised learning, compares and contrasts their performance in training a neural network to perform the task of classification on datasets. Both the techniques are found to be speedy and successful in training the neural networks. Further, their prediction accuracy is also found to be high. Except in the case of two datasets, the training and prediction accuracies of the Enhanced Fireworks Algorithm driven neural net are found to be superior to those of the Particle Swarm Optimization driven neural net.

1. INTRODUCTION

Most firms hold many data in their databases about the products and the customers buying those products. The problem with the data preserved into the databases is that they are hardly utilized to improve the business prospects. To improve their business prospects, firms have to make or sell the products with greater efficiency which means they have to know which products are preferred by the customers and which are not. However, it is not feasible to directly contact the customers to get some information about their needs and preferences. To understand their customers' needs, most of the firms have to analyze the databases that hold the customers' past purchase information.

In order to analyze these data sets and get valid, novel and potentially useful information, the data mining community has developed various analysis and data mining tools. These have been used to oper-

DOI: 10.4018/978-1-5225-0159-6.ch034

ate on a large amount of data to discover useful information, patterns and relationships to help make a decision. There are many kinds of data mining techniques depending on the dataset and the purpose of the study. In this paper, we will implement classification by supervised learning.

Classification is one of the most useful data mining techniques. It is a predictive data mining technique, making prediction about unknown data using known results found from different data (Bhardwaj, & Pal, 2011). The task of any supervised learner is typically to build a model that both fits a number of training examples (training data) and predicts the class into which new or unseen data falls, where the class of the instance may not be known (test data). Some of the well-known supervised learning algorithms include decision trees (Murthy, 1998; Gama, & Brazdil, 1999), probabilistic and statistical learning techniques, such as Naive Bayes classifiers (Bouckaert, 2004), Bayesian networks (Jensen, 1996; Cheng, & Greiner, 2001), support vector machines (Cristianini, & Shawe-Taylor, 2000), etc. A more detailed review of various classification techniques can be found in Kotsiantis, 2007).

The state-of-the-art technology to implement supervised learning is Artificial Neural Networks (ANN) (Carpenter, & Grossberg, 1992; Yamamoto, & Nikiforuk, 2000; Tang, Wang, Tamura, & Ishii, 2003; Khoshgoftaar, Van Hulse, & Napolitano, 2010). ANN consists of interconnected layers of neurons. Adjusting the connection weights so as to produce the correct output for a given input is called "learning". The conventional algorithm is the Back Propagation Algorithm (Du, Hou, & Li, 1992; Hsin, Li, Sun, & Sclabassi, 1995; Lee, Yang, & Ho, 2006) which systematically varies the weights of the connections by propagating the magnitude of error in the backward direction from the output to the input in the form of a feedback. However, this algorithm is found to be slow in convergence and is sensitive to the initial weights that are usually chosen randomly. Recently, Evolutionary Algorithms (EA) have been widely used in optimizing diverse engineering and business domain problems. The EA optimization being a meta-heuristic strategy can, in principle, be applied to any optimization problem since it does not exploit the problem domain heuristics. Moreover, being a population-based parallel search, it is robust and rapid in convergence. Recently, they have been used in training ANN (Abass, 2001; Xiao, Wu,Yang, 2001; Siu, Yang, Lee, & Ho, 2007).

Another area in Computational Intelligence that is gaining momentum is Swarm Intelligence (SI) (Bonabeau, & Theraulaz, 1999; Kennedy, Eberhart, & Shi, 2001; Gonsalves, 2014). Similar in structure and strategy to EA, SI performance is also robust and speedy. The Particle Swarm Optimization (PSO) algorithm based on the food-searching strategy of a swarm of birds or a school of fish is a well-known SI algorithm (Kennedy, & Eberhart, 1995; Kennedy, & Eberhart, 1997). Its applications in optimization are rapidly increasing (Krause, Cordeiro, Parpinelli, & Lopes, 2013; Olsson, 2011). This study demonstrates a novel approach of using PSO to train ANN to perform supervised learning.

Another recent SI optimization technique is the Fireworks Algorithm (FWA). It mimics the explosion of the fireworks in the night sky to achieve function optimization (Tan, & Zhu, 2010; Liu, Zheng, & Ying, 2014). Although it has been put to use in many applications it is not without deficiencies. One of the well-known shortcomings of FWA is that its performance deteriorates for functions whose optimum is far away from the origin. Several attempts have been made to improve the performance of the basic FWA leading to an Enhanced Fireworks Algorithm (EFWA) (Pei, Zheng, Tan, & Takagi, 2012; Jianhua, Zheng, Ying, 2013; Zheng, Janecek, & Tan, 2013; Zhang, Zhang, & Zheng, 2014). In this study, we demonstrate how to use the EFWA to train ANN for supervised learning. Finally, we compare and contrast the two SI techniques in training ANN. The experimental results show that the two techniques are successful in training the neural nets to perform classification.