Learning Fuzzy Network Using Sequence Bound Global Particle Swarm Optimizer

Satchidananda Dehuri, Fakir Mohan University, India
Sung-Bae Cho, Yonsei University, Korea

ABSTRACT
This paper proposes an algorithm for classification by learning fuzzy network with a sequence bound global particle swarm optimizer. The aim of this work can be achieved in two folded. Fold one provides an explicit mapping of an input features from original domain to fuzzy domain with a multiple fuzzy sets and the second fold discusses the novel sequence bound global particle swarm optimizer for evolution of optimal set of connection weights between hidden layer and output layer of the fuzzy network. The novel sequence bound global particle swarm optimizer can solve the problem of premature convergence when learning the fuzzy network plagued with many local optimal solutions. Unlike multi-layer perceptron with many hidden layers it has only single hidden layer. The output layer of this network contains one neuron. This network advocates a simple and understandable architecture for classification. The experimental studies show that the classification accuracy of the proposed algorithm is promising and superior to other alternatives such as multi-layer perceptron and radial basis function network.

Keywords: Classification, Fuzzy Network, Multi-Layer Perceptron, Particle Swarm Optimization, Radial Basis Function Network

INTRODUCTION
In the past few decades, information technology and the World Wide Web (WWW) have created stacks of innovations in the area of marketing style of companies. More businesses and companies are collecting highly informative and valuable data in a large scale. The huge amount of data can be a gold mine for business management and marketing. It is therefore increasingly important to analyze the data. However, timely and accurately processing tremendous volume of data with traditional methods (Michie et al., 1994) is a difficult task. For example, using multi-layer perceptron (MLP) in data mining is not likely produce any useful results (Edelstein, 1996), because it does not have a clean interpretation of the model and its longer training time can make frustration. The ability to analyze and utilize massive data lags far behind the capability of gathering and storing it. This gives rise to new challenges for businesses and researchers in the extraction of useful information.

DOI: 10.4018/ijfsa.2012010104
Data mining—a core step of knowledge discovery in databases (Piatetsky-shapiro et al., 1996; Han & Kamber, 2001), is defined as a process of employing one or more computer learning techniques to automatically analyze and extract knowledge from vast amount of data contained within a database. The purpose of data mining is to identify trends and patterns in data. Classification (Duda et al., 2001) is one of the widely used techniques of data mining. In addition to classification, there are many important tasks such as association rule mining (Agrawal, 1993), clustering (Jains & Dubes, 1988), regression analysis (Chen, 2011), summarization, etc. in the area of data mining. However, classification is a fundamental activity in pattern recognition (Theodoridis & Koutroumbas, 2006; Bishop, 2006), data mining and so forth. Given predetermined disjoint target classes \( \{C_1, C_2, \ldots, C_m\} \), a set of input features \( \{F_1, F_2, \ldots, F_n\} \) and a set of training data \( T \) with each instance taking the form \( <a_1, a_2, \ldots, a_m> \), where \( a_i = \{1,2, \ldots, m\} \) is in the domain of attribute \( F_i \), \( i=1,2, \ldots, m \) and associated with a unique target class label the task is to build a model that can be used to predict the target category for new unseen data given its input attribute values.

There are many classifiers like naïve Bayes (Robert, 2001; Winkler, 2003; Gelman, et al., 1995), linear discriminant (McLachlan, 2004), k-nearest neighbour (Dasarathy, 1990; Herbrich, 2001), MLP (Bishop, 1995) and its variant, decision tree (Breiman et al., 1984; Quinnlan, 1992), and many more are commonly available in the specialized literatures. We can use the existing techniques of pattern recognition for classification in the context of data mining but these algorithms were designed only for small dataset. Therefore, either we can prefer to design a new algorithm or we can reengineer the existing classification algorithms of pattern recognition such that a gamut of data can handle efficiently. Data mining does not compete with traditional methods. However, it offers better solutions in certain classes problems than traditional methods. Data mining methods and algorithms extract useful regularities from large data archives, either directly in the form of knowledge or indirectly as functions that allow predicting, classifying or representing regularities in the distribution of the data.

A neural network classifier like MLP is a parallel computing system of several interconnected processor nodes. The strength of MLP is the ability to construct non-linear boundary with high classification accuracy. However, the main weakness of this network lies in its architectural complexity and training speed. It requires many passes to build. This means that creating the most accurate models can be computationally expensive. Craven and Shavlik (1997) was described a neural network learning algorithm that can suitably address the issues of model comprehensiveness and training speed. However, it leaves several rooms for carrying out more research on noisy and uncertain data.

The paper by Bellman et al. (1996) was the starting point in the application of fuzzy set theory to classification. Since then, researchers have found several ways to apply this theory to generalize the existing classification methods, as well as to develop new algorithms. There are two main categories of fuzzy classifiers: fuzzy if-then rule-based and non if-then rule fuzzy classifiers. The second group may be divided into fuzzy k-nearest neighbours and generalized nearest prototype classifiers (GNPC). Several approaches (Nauck & Kruse, 1997; Uebele et al., 1995; Chakraborty & Pal, 2004; Abe & Thawonmas, 1997; Roubos et al., 2003) have been proposed for automatically generating fuzzy if-then rules and tuning parameters of membership functions for numerical data. These methods fall into three categories: neural-network-based methods with high learning abilities, genetic (evolution)-based methods with the Michigan and Pittsburg approaches, and clustering-based methods. There are several methods (e.g., Zogala, 2000; Ishibuchi et al., 2001; Zhou & Khotanzad, 2007) that combine the categories that have proved as effective in improving classification performance.

Recently, a new direction in the fuzzy classifier design field has emerged: a combination of multiple classifiers using fuzzy sets (Kuncheva,
Introduction to Smart Manufacturing: Value Chain Perspective for Innovation and Transformation
ZongWei Luo (2014). *Smart Manufacturing Innovation and Transformation: Interconnection and Intelligence* (pp. 1-19).
www.igi-global.com/chapter/introduction-to-smart-manufacturing/102098?camid=4v1a