Chapter 9
Evaluation of Parameter Settings for Training Neural Networks Using Backpropagation Algorithms:
A Study With Clinical Datasets

Leema N.
Anna University, Chennai, India

Khanna H. Nehemiah
Anna University, Chennai, India

Elgin Christo V. R.
Anna University, Chennai, India

Kannan A.
Anna University, Chennai, India

ABSTRACT

Artificial neural networks (ANN) are widely used for classification, and the training algorithm commonly used is the backpropagation (BP) algorithm. The major bottleneck faced in the backpropagation neural network training is in fixing the appropriate values for network parameters. The network parameters are initial weights, biases, activation function, number of hidden layers and the number of neurons per hidden layer, number of training epochs, learning rate, minimum error, and momentum term for the classification task. The objective of this work is to investigate the performance of 12 different BP algorithms with the impact of variations in network parameter values for the neural network training. The algorithms were evaluated with different training and testing samples taken from the three benchmark clinical datasets, namely, Pima Indian Diabetes (PID), Hepatitis, and Wisconsin Breast Cancer (WBC) dataset obtained from the University of California Irvine (UCI) machine learning repository.

DOI: 10.4018/978-1-6684-2408-7.ch009

INTRODUCTION

Researchers aim to build a computing system that will operate intelligently like a human brain. The Artificial Neural Network (ANN) facilitates the information processing in an intelligent manner (Akinyokun, 2002; Bezdek, 1993), and is inspired by the biological neural system. A biological nervous system is a large interconnection of neurons located within the brain. The functional equivalent of an artificial neuron is known as computational neuron or a node (Eluyode, Akomolafe & MNCS, 2013). These neurons are structured hierarchically by layers and interconnected between them like the biological nervous systems. Artificial neural network determines the rate of adjustment required for internal network parameters. This adjustment is known as learning or training the network. The neuron functions are described by the activation function. Activation functions are used in the hidden and the output layer. Hidden layer implements the non-linear activation function, whereas the output layer implements the linear activation function. Linear activation function used in neural network training is purelin and the non-linear activation functions are hardlim, sigmoid and logistic (Sharma, 2014).

Parameters of both biological and Artificial Neural Networks are structures, layers, number of neurons, the functional capabilities of neurons, their learning capabilities, processing elements, connections, strength, processing speed, style of computation, information storage, signal transduction, information transmission communication media selection and fault tolerance. Major difficulty faced in correlating artificial neural networks with biological neural networks are adjusting weights and synaptic strengths. Weights are altered mathematically in an ANN, based on differences in error values. Synaptic strengths are modified in response to synaptic activity. A simple feed-forward system behaves similar to biological neurons and they are used for pattern recognition. Once input values are given to the input layer, neuron computes the output, layer by layer. The dependence of output values and input values require adjusting every weight, and threshold, which can be complex and time consuming. After training is complete, the network is able to give reasonable outputs for any type of input, even if the test data does not match with the training data. This is referred to as the generalization capability of the network. In that case, the ANN attempts to determine the best output depending on its training method.

Based on the structure, ANN is divided into two types, namely, single layer neural network and multilayer neural network. Single layer neural network is used for linearly separable problems, whereas the multilayer neural network is used for linearly non separable problems. One major drawback of a single layer network is that it can predict the output, which is similar to the input pattern. For many practical problems, very similar input patterns may have very different output requirements (FFNN, 2010). To overcome the above limitation multilayer neural network has been developed with one or more hidden layers, called multi-layer perceptron (MLP) networks. Hidden layer in the MLP is used to deal with nonlinear relationships between input features and the output layer is used to obtain the predicted output. This MLP can be used to solve many real world problems like predicting the future trends based on the historical data (Kosko, 1994.). ANN have been implemented in many science and engineering fields such as, decision making and control, biological modeling, health care and medicine, marketing, engineering and manufacturing for classification task (Krasnopolsky & Dé ricChevallier, 2003; Coppin, 2004; Basheer & Hajmeer, 20007; He, Wu & Gong, 1992).

In machine learning, Backpropagation (BP) is a supervised learning algorithm for training the Artificial Neural Network (ANN). Most of the researchers used different BP algorithms to train ANN without knowing the performance of different BP algorithms and network parameter adjustments. Backpropagation training algorithms receive the inputs, adjust the weights and produce the required output.