International Conference Mathematical and Computational Biology 2011 International Journal of Modern Physics: Conference Series

Vol. 9 (2012) 448-455

☐ World Scientific Publishing Company DOI: 10.1142/S2010194512005533



LOCAL MINIMA PROBLEM IN BACK PROPAGATION ALGORITHM ON CLASSIFICATION PROBLEMS

NORHAMREEZA ABDUL HAMID

Post Graduate Centre, Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia gi090007@siswa.uthm.edu.my

NAZRI MOHD NAWI

Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia nazri@uthm.edu.my

ROZAIDA GHAZALI

Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia rozaida@uthm.edu.my

MOHD NAJIB MOHD SALLEH

Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia najib@uthm.edu.my

This paper presents a new method to improve back propagation algorithm from getting stuck with local minima problem and slow convergence speeds which caused by neuron saturation in the hidden layer. In this proposed algorithm, each training pattern has its own activation functions of neurons in the hidden layer that are adjusted by the adaptation of gain parameters together with adaptive momentum and learning rate value during the learning process. The efficiency of the proposed algorithm is compared with the conventional back propagation gradient descent and the current working back propagation gradient descent with adaptive gain by means of simulation on three benchmark problems namely iris, glass and thyroid.

Keywords: back propagation; gain; momentum; learning rate.

1. Introduction

Artificial Neural Network (ANN) is computational models whose architecture and operation is inspired by human knowledge of biological nervous system. Due to its ability to solve some problems with relative ease of use, robustness to noisily input data, execution speed and can analyse complicated systems without accurate modeling in