Measuring the Learnability of a Parameter in Multi-Layer Perceptron Neural Network

Siddhaling Urolagin\textsuperscript{a}, Prema K V\textsuperscript{b} and N V Subba Reddy\textsuperscript{c}

\textsuperscript{a}Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal, India, Contact: siddeshu@ yahoo.com

\textsuperscript{b}Mody Institute of Technology and Science, Rajasthan, India, Contact: drprema.mits@gmail.com

\textsuperscript{c}Mody Institute of Technology and Science, Rajasthan, India, Contact: drnvsreddy@rediffmail.com

During the training process of Multi-Layer Perceptron (MLP) not all parameters will learn in the equal proportion; some are more discriminative than others. How network parameters are learning during the training process? The objective analysis of this question has many applications in the neural network, such as network freezing, network pruning, fault tolerance etc.. A method is developed in this research to measure the learning capability of a parameter called as learnability. The learnability is defined as the contribution of that parameter to reduce the network error as it undergoes training. The association between surface of performance index and learnability is studied. On the steeper regions, the learnability is higher than on the flatter regions of the surface. Further studies are carried out by removing a parameter based on learnability from a trained network. On removal of a parameter with higher learnability has drastically affected the Mean Square Error (MSE) and classification rate. On the other hand, a parameter with a lower learnability has least impact on MSE and classification rate of a trained network. In this research, the learnability measurement is applied on network freezing and network pruning to improve the performance of MLP network.

Keywords: Freezing, Learnability, Multi-Layer Perceptron, Pruning.

1. INTRODUCTION

Artificial Neural Network (ANN) is a biologically inspired model. Among the ANNs the backpropagation neural network (also called a Multi-Layer Perceptron (MLP)) is the most widely used network. There are many applications of MLP in the area of pattern recognition, computer vision, speech recognition and speech synthesis [1], etc.. To name a few, handwriting recognition [2][3], speech recognition [4][5], product inspection[6], optimization of computation [7], forecasting [8], fault detection [9], medical diagnosis [10][11] etc.. The MLP consist of a number of highly interconnected computing units, called neurons or nodes. Each unit receives inputs from other units in the network, or from the outside world and calculates an output based on these inputs. Associated with each connection (sometimes called a synapse) between the units is a weight. Each neuron has a bias associated with it. Typically, an architecture is chosen and the network undergoes training process to reduce the network error or performance index. As training progresses from initial weight and bias values, the network parameters such as weights and biases are updated according to the learning algorithm to reduce the performance index [12]. Not all the network parameters equally learn the input-output mapping. Some parameters would hold more discriminating capability while others are not so effective. It is interesting to investigate how the network parameters are
the input-output mapping, some hold higher discrimination than others. In this research a method is developed to measure the learning capability (called as learnability) of a network parameter. Investigations are carried-out on the association of learnability and performance index’s surface. It is found that learnability factors of parameters on the steeper region are higher than on the flatter region. Further, the behavior of learnability of several parameters during the training process is analyzed on MNIST handwritten database. In addition, the effect of removing a parameter based on learnability from a trained network is studied. It is observed that the parameter having higher learnability has the more impact on MSE and classification rate of a network. In this research, learnability measurement has been applied on two important problems: network freezing and network pruning. The network freezing is one of the methods to improve the training time. It involves identifying the parameters those learnability is low and freezing of these parameters from the training. Once parameters are frozen, they will not participate in the training process. The network pruning is the method to improve the generalization performance by eliminating least significant parameters based on the learnability. Several experiments are conducted on feedforward neural network trained with backpropagation learning algorithm on MNIST handwritten numeral database. The network freezing and pruning results are compared with standard feedforward neural network where no freezing or no pruning is performed. The experimental results show better performance in terms of reduction in training times and improvement in classification rates and mean squared errors.

REFERENCES

Measuring the Learnability of a Parameter in Multi-Layer Perceptron Neural Network


Dr. Siddhaling Urolagin has completed B.E in 2001 from Goa Institute of Technology, Belgaum, M.Tech in 2004 and Ph.D in 2010 from Manipal Institute of Technology, Manipal, all in Computer Science and Engineering. He is currently working as Associate Professor in Department of Computer Science and Engineering at MIT, Manipal. His research interests include Gabor Filter based Feature Extraction, Artificial Neural Networks and Document Image Analysis. He is also carried out an AICTE funded project. He published several international/national journal and conference papers.
Dr. Prema K V, B.E, M.E, Ph.D. She is currently the Professor at Department of CSE, Faculty of Engineering and Technology, MITS University, Lakshmangarh. Her research interests include Artificial Intelligence, Pattern Recognition, Neural Networks, Computer Networks and Network Security. She has published around 25 papers in reputed journals/conferences. She is currently guiding 3 research scholars. She is also the principle investigator for one of the AICTE projects entitled “Design of an Intelligent System for Pattern Recognition using Soft Computing Models”.

Dr. N V Subba Reddy received Master of Engineering Degree, in Computer Science and Engineering, and PhD in Computer Science and Engineering from Indian Institute of Science, Bangalore and University of Mysore respectively. He is currently working as Vice Chancellor of Mody Institute of Technology and Science, Rajasthan. He was working as professor in Computer Science and Engineering Department and Associate Director Research and Development, at Manipal Institute of Technology, Manipal University. He was also working as Director Sikkim Manipal Institute of Technology at Sikkim Manipal University, Sikkim. He has guided several Ph.D’s in the area of Pattern Recognition, Neural Networks, Fuzzy Logic and Genetic Algorithms. Currently he is guiding Ph.D’s in Document Image Processing, Character Recognition, and Machine Intelligence and Computer Cognition.