# Classification of Data Using Multilayered Perceptron Neural Network

**Ariffuddin Joret[1], Siti Zuraidah Zainudin[1], Nor Ashidi Mat Isa[2], Jiwa Abdullah[1], Kamal Zuhairi Zamli[2], Muhammad Faiz Liew Abdullah[1], Asmarashid Ponniran[1]**

[1]Department of Electrical Engineering, Universiti Tun Hussein Onn Malaysia

[2]School of Electrical and Electronic Engineering, Universiti Sains Malaysia

**Abstract:** Neural networks become widely in its application to help human in the identification process nowadays. However, to have a neural network system which is able to provide 100% accuracy performance and providing optimum value structure is difficult. In this project, multilayer perceptron neural network with Backpropagation, as its algorithm has been trained and tested for its accuracy. The multilayer perceptron neural network has been tested using gravel data which have been categorized into 6 categories, and gives 82.90% of the network classification performances for optimum structure with 16 hidden node.

**Keywords :** Neural network, Multilayered perceptron, Beackpropagation learning algorithm

## 1. Introduction

Neural network can be imagine as human brain function. In human brain, it have billions of neuron cells which connected and communicated each other and works to think and do smart think. Many computer scientist think to make a computer based on human brain function idea and this is what we called neural network. Neural network is a one of modelling method that had been used to solve any complex problems. Neural network processing is done by looking at the data in the training set and then learn the pattern and next is to create the ability to classify new patterns and make predictions. The learning process is accomplished by adjusting the weights which connect between the neurons. According to Paolo Antogmettim[16], neural network is a modelling of dynamic parallel computing system, that contains a lot of processing units, where each unit representing the biological characteristics of brain's neurons that are connected to one another in a particular configuration, which can generate useful information depends on the input value received. Neural networks are very popular applications in fields such as oil and gas as used by Heydari et al. Al, [6], and in transportation field as used by Kurokawa and Takeshita,[10]. Furthermore, other research and application of neural networks is in the power system by M. Tarafdar Haque, and A.M. Kashtiban,[5] also in the field of data classification as done by Pardo et al.[17] and Schmitz & Aldrich.[19]. The neural networks have been extensively used in medicine and data classification field. The major strengths of neural network are having capability to adapt, improve and optimize themselves according to variety and change of input information. These capabilities had help to improve the efficiency and reliability of inspection and diagnosis of disease from medical images. Veropoulos et al.[21] had used image processing technique and neural network to detect the presence TB Tuberculosis bacilli in auramine - stained sputum specimens. Technically, neural network can be divided into three main layers which are input layer, hidden layer and output layer. This layer consists of simple processing elements called as nodes. Each node connected to each other with the weights. The neural network system consists of two main types of data sets known as training data and testing data. However, having a neural network system that is capable of providing optimum value structure is

difficult. It is hard to have a neural network system that can give a 100% performance of accuracy. According to Masters, T. [11], and Remus, W. & O'Connor, M. [18], one hidden layer can provide better performance in neural networks. Most of the problems predicted by the neural network involving the use of the second hidden layer will not have the greatest impact on the accuracy and performance of the network model. Moreover, by adding second hidden layer will only slow down the convergence due to additional neurons, weights, coefficients, and thus will slow down the work and the training data without impacting the efficiency of the network model. Based on this, the project had been designed to use one hidden layer of neural network structure. The main objective for this project is to determine the optimal structure of MLP neural network by testing its capability and accuracy using scaled conjugate gradient back propagation algorithm in gravel data. In the project, to determine the optimal network structure, number of hidden nodes and the overall performance of the tested data will been analyzed and recorded in the table. Data will been tested starting with hidden nodes 1 until 100.

## 2. Neural Network

Neural network gained its knowledge from past experiences to deal with a problem and situation. Neural network take a solution before to build up the system in decision making neurons. Neural network which consist of three layers of neurons will learn through the adjustment between the weights and the associated layer. The answer to this network was compared with that required answer over and over during the training phase. If there is a problem, then the system will learn the problem thus the sets of weights will be adjusted until it can produces the best output result.

### 2.1 Multilayered Perceptron neural network (MLP)

In theory, MLP network is a non-linear neural network system which consists of a set of one input layer, one or more hidden layer and output layer.

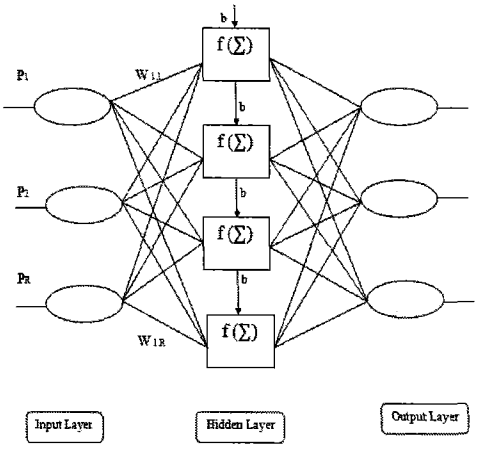



**Figure 1: Basic architecture of MLP neural networks**

Figure 1 shows the basic MLP neuron network structure with R as the number of inputs. Each input will be adjusted by the weights, w. Total input weights, w and bias, b would lead to a transfer function, f. Bias, b is approximately the same as the weights, w but the bias has a value 1 as constant value. Weight value, w and bias, b is a scale parameter which can be modified. Usually, the MLP network will use log-sigmoid transfer function as shown in Figure 2 in the hidden layer. In most neural network applications, log-sigmoid transfer function is often used in object recognition. Transfer function will then calculate the value of output from the net input. This operation of log-sigmoid transfer function (log-sig) will produce an output between 0 and 1 which indicates the value of the net input from negative to positive infinity.

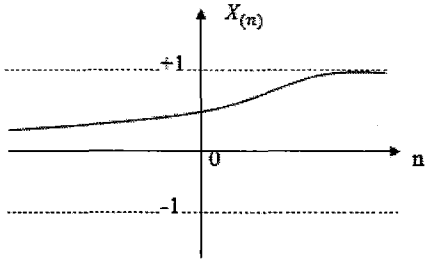



**Figure 2 : Log-Sigmoid Transfer Function**

In process of analysing the neural networks, data can be divided into two sets of data that known as the training and testing data sets. However, to develope the neural network system, usually another subset of data needed known as validation

data set. Validation set is often used as a confirmation data of neural network that been analyzed by stopping the analysis when the chosen validation data has been exceeded. In the training phase, the weights, w and bias, b will be modified over and over again to obtain the optimal network structure. While in the testing phased, the weight,w and bias,b will be fixed and used to calculate the output.

## 2.2 Backpropagation, BP

To determine the optimal network structure, the MLP network can be trained by using backpropagation, BP algorithm. This BP technique will determine, whether the gradient of the network performance with network weights, w or by determining the Jacobian on the error in the network with the weights, w. The simplest implementation of BP learning updates the network weights and biases in the direction in which the performance function decreases most rapidly, which are in the negative of the gradient.

One iteration of this algorithm can be written as

$$x_{k-1} = x_k - \alpha_k g_k \qquad (1)$$

where;

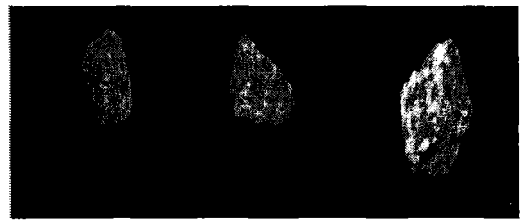
$x_k$ = vector of current weights and biases

$g_k$ = current gradient

$\alpha_k$ = leraning rate

## 3. Simulation

MLP neural network will be trained by using scaled conjugate gradient backpropagation. MLP neural network goes through two phases which is training phase and testing phase. In this project, gravel data have been used. The Gravel data which can be divided into six categories had been extracted for its 7 features based on three sides of view of its image. The three side of view of the gravel is the above view, $s_1$, the front view, $s_2$, and the side view, $s_3$ as shown in figure 3.



(a) (b) (c)

**Figure 3 : Gravel image; (a) front view, (b) above view and (c) side view.**

In this project, the distribution of data can be divided into ;

1. Training = 70%
2. Validation = 5%
3. Test = 25%

This means that the total number of data used to determine the optimal structure for MLP neural network of the gravel data is about 590 samples. 412 data samples are used as training set, 30 samples as the validation set and 148 samples as the testing sets.

## 4. Result and Analysis

Based on this project, the main objective was to test the accuracy of the performance of the MLP neural network that used the scaled conjugate gradient backpropagation as a training algorithm. In addition, by carrying out this project, the optimal structure for MLP neural network can be determined. Analysis is performed starting from the value of hidden nodes of 1 until 100. Figure 4 shows the result of performances for the MLP neural network of the gravel data with the hidden node.

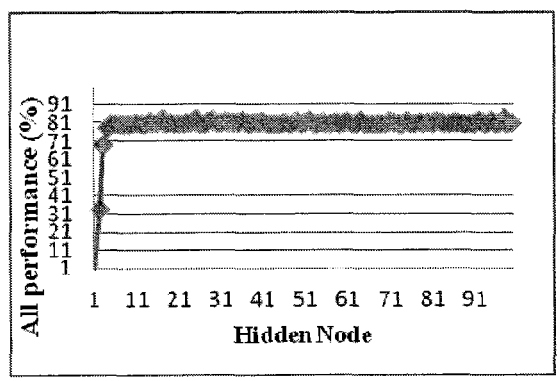



**Figure 4: Graph Overall performance vs hidden node**

| Hidden node | Overall performance (%) |
|---|---|
| 1 | 33.20% |
| 16 | 82.90% |
| 97 | 83.40% |

**Table 1**

Based on Table 1, one hidden node of MLP neural network gives 33.20% of overall performance which also mean the minimum performance of the system. By having 97 hidden node in MLP neural network in classification of gravel data, the reading show 83.40% for overall performance and thus state the highest performance for the system. The optimal structure for the MLP network is shown at hidden node of 16 with 82.90% of overall performance.

## 5. Conclusion

In this project, a system of multilayered perceptron neural network (MLP) had been tested its accuracy in indentifying gravel data. This system had been trained by using scaled conjugate gradient backpropagation algorithm. This system achieved a maximum performance with 83.40% using 97 hidden nodes. Meanwhile, the optimal structure for the MLP neural network system in classifying the gravel data is achieved at 16 hidden nodes with 82.90% of performances

## References


[1] hens. Cowan CFN. Billings SA, Grant PM. Parallel RecursivePrediction Error Algorithm for training Layered Neural Networks. International Journal of control, 1990 ; 51 (6); 1215-28

[2] hen ST, Hsiao YH, Huang YL, Kuo SJ, Tseng HS, Wu HK, et al. Comparative Analysis of Logistic Regression, Support Vector Machine and Artificial Neural Network for the Differential Diagnosis of Benign and Malignant Solid Breast Tumors by the Use of Three-Dimensional Power Doppler Imaging. Korean Journal of Radiology. 2009;10(5):464.

[3] . Bhattacharjee, M. K. Bhowmik, M. Nasipuri, D. K. Basu & M. Kundu, A Parallel Framework for Multilayer Perceptron for Human Face Recognition, International Journal of Computer Science and Security (IJCSS), Volume (3): Issue (6)

[4] avrilov, A.V. Hybrid Neural Network Based on Models of Multi-Layer Perceptron and Adaptive Resonance Theory. In: Proceedings of 9th Korean-Russian International Symposium on Science and Technology KORUS-2005, NSTU, Novosibirsk (2005) 119-122.

[5] aque.M.T. and Kashiban. A.M. Application of neural networks in power system; A review. World Academy of Science, Engineering and Technology, University of Tabriz, 2005.pp 53-57

[6] eydari, E., 2000. Porosity loss, fluid flow, and mass transfer in limestone reservoirs: application to the Upper Jurassic Smackover Formation, Mississippi. American Association of Petroleum Geologists Bulletin 84 (1), 100– 118.

[7] EEE Transaction on Automatic Control, Vol. AC-30, No.9 sept 1985. Theory and practise of recursive identification- Lennart Ljung and Torsten Soderstrom ( Cambridge, MA: M.I.T. Press, 1983) Reviewed by P.R.Kumar

[8] uan. J.M.M, Alfonso. P.P, & Pilar. M.G, Artificial neural networks applied to forecasting time series. Facultad de Psicología Universidad de las Islas Baleares, 2011

[9] . J. HOPFIELD Neural networks and physical systems with emergent collectivecomputational abilities. Proc. NatL Acad. Sci. USA Vol. 79, pp. 2554-2558, April 1982 Biophysics

[10] urokawa, T. and Takeshita K. (2004), "Air transportation planning using neural networks as an example of the transportation squadron in the Japan Air self-defense force", *Systems and Computers in Japan*, 35(12), 1223-1232.

[11] asters, T. (1993). *Practical Neural Network Recipes in C++*, San Diego: Academic Press.

[12] at Isa N.A, Subramanium.E, Mashor M.Y & Osman N.H, Fine needle aspiration cytology evaluation for classifying breast cancer using artificial neural network, AM.J applied sci, 4(12) ; 999-1008, 2007

[13] M
ashor M.Y Hybrid Multilayered perceptron networks. Internasional journal of system science.2000; 31 (6) ;771-85

[14] O
thman, N.H. 1995. Pap Smears: A kelantan experience; Is It An effective method of cervical cancers screening? A study of 2229 Cases. *Malaysia Journal of pathology* 17 (1)

[15] O
sman, M.K. Mashor, M.Y. & Jaafar, H. Improving performance of hybrid multilayered perceptron network for classification of mycrobacterium tuberculosis in tissue using modified recursive prediction error-extreme learning machine (2011). Vol. 1, No.3, pp.57-65

[16] P
aolo Antogmettim, Neural Networks Concepts, Application, and Implementations Vol II, Prentice Hall Advances Reference Series Engineering

[17] P
ardo, M., Sberveglieri. G., Gardini. S., and Dalcanale. E. (2000). A. Hierarchical Classification Scheme for An Electronic Nose. Sensor and Actuators Vol. B. 69. pp.259-365

[18] R
emus, W. & O'Connor, M. (2001). Neural networks for time series forecasting. In J. S. Armstrong (Ed.), *Principles of Forecasting: A Handbook for Researchers and Practitioners* (pp. 245-256). Norwell, MA: Kluwer Academic Publishers.

[19] S
chmitz, G.P.J, Aldrich.c. ANN-DT: An algorithm for extraction of decision trees from artificial neural networks, IEEE Transactions on Neural Networks, 10(6), 1392-1401

[20] S
. Bhama and H. Singh, "Single layer neural networks for linear system identification using gradient descent technique," *IEEE Trans. Neural Networks*, vol. 4, pp. 884–888, Sept. 1993.

[21] V
eropoulos K, Campbell C, Learmonth G, knight B, Simpson J. The Automed identification of tubercle bacilli using image processing and neural computing technique. 8th international conference on artificial neural network, ; 2-4 september skovde, sweden.; Citeseer; 1998.p.797-802.