

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320-088X



IJCSMC, Vol. 3, Issue. 11, November 2014, pg.455 – 464

REVIEW ARTICLE

Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural Network

Foram S. Panchal¹, Mahesh Panchal²

¹Department of Computer Science & Engineering,
Kalol Institute of Technology, Gujarat, India

²Department of Computer Science & Engineering,
Kalol Institute of Technology, Gujarat, India

¹ forampanchal187@gmail.com

² mkhpanchal@gmail.com

Abstract - Artificial Neural Networks are most effective and appropriate for pattern recognition and many other real world problems like signal processing, Classification problems. Superior results in pattern recognition can be directly provided in the forecasting, classification and data analysis. To bring proper results, ANN requires correct data pre-processing, architecture selection and network training but still the performance of a neural network depends on the size of network. Selection of hidden neurons in neural network is one of the major problems in the field of Artificial Neural Network. The random selections of hidden neurons may cause the problem of either Underfitting or Overfitting. Overfitting arises because the network matches the data so closely as to lose its generalization ability over the test data. The proposed method finds the near to optimal number of hidden nodes after training the ANN from real world data. The advantage of proposed method is that it is not approximately calculating number of hidden nodes but based on similarity between input data, they are calculated.

Keywords – Artificial Neural Network, Supervised learning, Learning rate, Hidden nodes, Training samples

1. INTRODUCTION

Neural Networks are most effective and appropriate for pattern recognition and many other real world problems like signal processing, Classification problems. Superior results in pattern recognition can be directly provided in the forecasting , classification and data analysis. To bring proper results, NN require correct data pre-processing, architecture selection and network training but still the performance of a neural network depends on the size of network. Selection of hidden neurons using the neural networks is one of the major problems in the field of Artificial Neural Network. Sometimes an overtraining issue is exists in the design of NN training process. Over training is same to the issue of overfitting data. Overtraining arises because the network matches the data so closely as to lose its generalization ability over the test data. This paper gives some introduction to artificial neural network and its activation function. Also give information about learning methods of ANN as well as various application of ANN. In this paper we introduce information regarding effect of Underfitting and overfitting and approaches to find hidden nodes in hidden layer. We also done some experiments on sale-forecasting data and check MSE for that and give the conclusion.

2. ARTIFICIAL NEURAL NETWORK

Artificial Neural Network is an information processing system which is inspired by the models of biological neural network [1]. It is an adaptive system that changes its structure or internal information that flows through the network during the training time [2]. In terms of definition Artificial Neural Network is Computer simulation of a "brain like" system of interconnected processing units.

2.1 Architecture

Artificial neural network (ANN) is a machine learning approach that models human brain and consists of a number of artificial neurons.[2] Neuron in ANNs tend to have fewer connections than biological neurons. Each neuron in ANN receives a number of inputs. An activation function is applied to these inputs which results output value of the neuron. Knowledge about the learning task is given in the form of examples called training examples.

Architecture of the Artificial Neural Network consists Input layer, Hidden layer, Output layer.

2.1.1 Input Layer :

The Input Layer is a layer which communicates with the external environment. Input layer presents a pattern to neural network. Once a pattern is presented to the input layer, the output layer will produce another pattern. It also represents the condition for which purpose we are training the neural network [3].

2.1.2 Output Layer :

The Output layer of the neural network is what actually presents a pattern to the external environment. The number of output neurons should be directly related to the type of the work that the neural network is to perform [3].

2.1.3 Hidden Layer :

The Hidden layer of the neural network is the intermediate layer between Input and Output layer. Activation function applies on hidden layer if it is available. Hidden layer consists hidden nodes. Hidden nodes or hidden neurons are the neurons that are neither in the input layer nor the output layer [3].

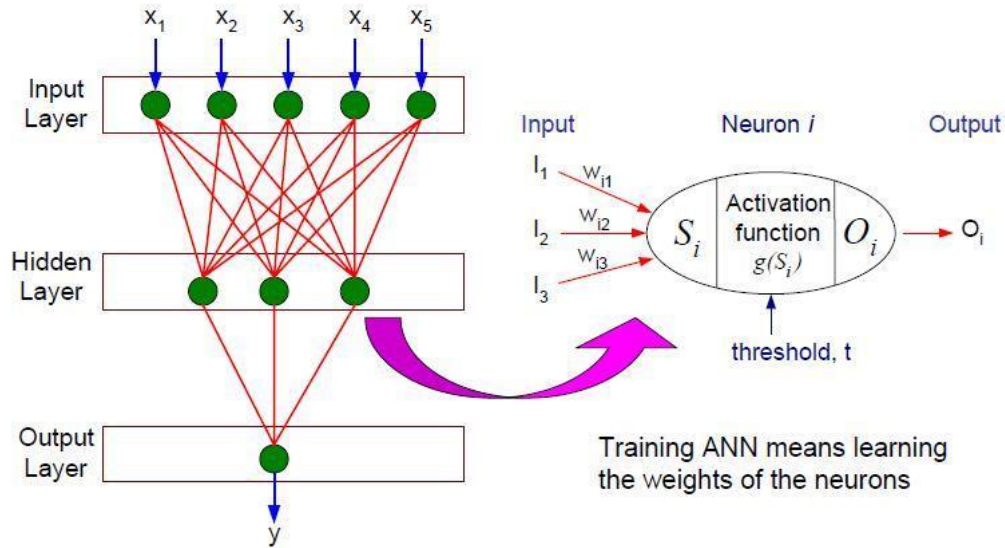


Fig.1 Architecture of ANN

2.1.4 Types of neural network:

There are several types of neural networks: Feed forward NN, Radial Basis Function (RBF) NN, and Recurrent neural network [4].

2.1.4.1 Feed forward neural network

A feed forward neural network begins with an input layer. This input layer is connected to a hidden layer. This hidden layer is connected with other hidden layer if any other one is there otherwise it is connected to the output layer. In common most of the neural networks have at least one hidden layer and it is very rare to have more than two hidden layer [10].

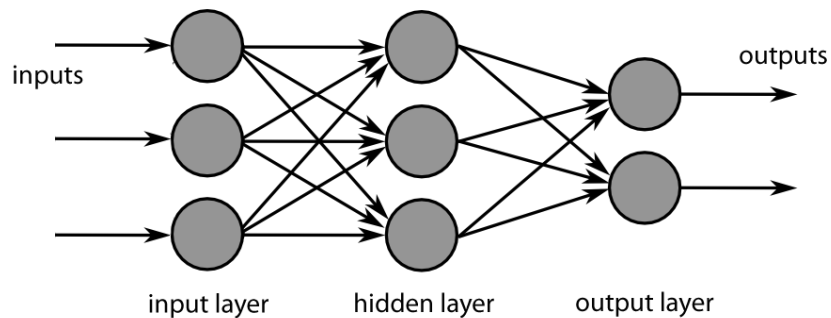


Fig.2 Feed forward neural network

2.1.4.2 Radial basis function network

Radial basis function (RBF) networks are feed-forward networks trained using a supervised training algorithm. They are typically configured with a single hidden layer of units whose activation function is selected from a class of functions called basis functions. While similar to back propagation in many respects, radial basis function networks have several advantages. They usually train much faster than back propagation networks [11].

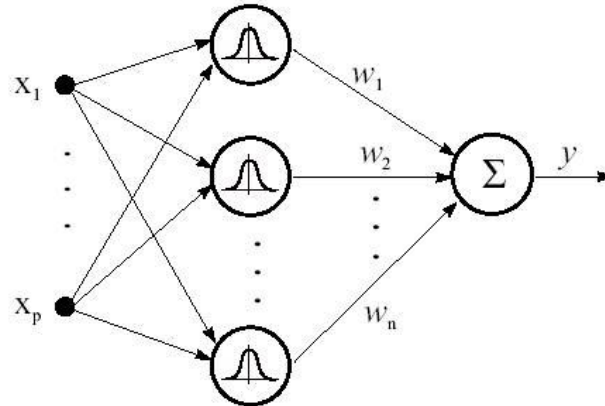


Fig.3 Radial basis function neural network

2.1.4.3 Recurrent neural network

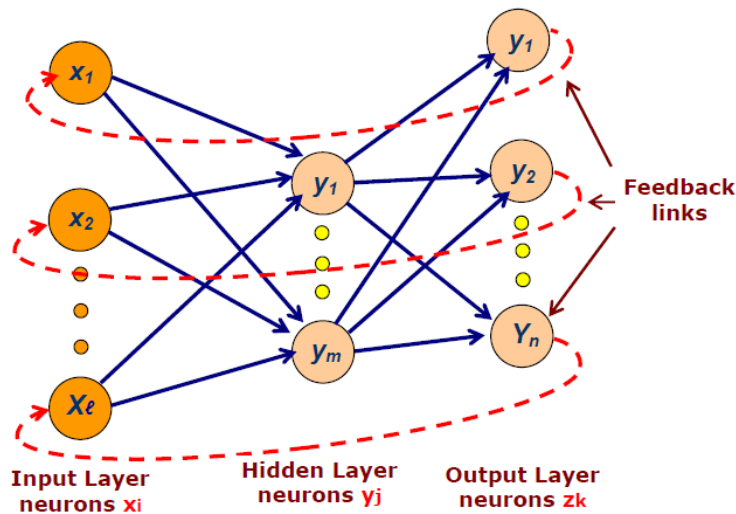


Fig.4 Recurrent neural network

The fundamental feature of a recurrent neural network is that the network contains at least one feedback connection. There could also be neurons with the self-feedback links it means the output of a neuron is feedback into itself as input [12].

2.2 Application of ANN

There are various types of applications of Artificial Neural Network.

- System identification and control
- Game-playing and decision making (backgammon, chess, poker)
- Pattern recognition (radar systems, face identification, object recognition and more)
- Sequence recognition (gesture, speech, handwritten text recognition)
- Medical diagnosis
- Financial applications (e.g. automated trading systems)
- Data mining (or knowledge discovery in databases, "KDD")
- Visualization and e-mail spam filtering

2.3 Learning Methods of ANN

The learning methods in neural networks are classified into three basic types [5] :

- 1) Supervised Learning
- 2) Unsupervised Learning
- 3) Reinforced Learning

In Supervised learning a teacher is present during learning process. In this process expected output is already presented to the network. Also every point is used to train the network.

In Unsupervised learning a teacher is absent during the learning process. The desired or expected output is not presented to the network. The system learns of its own by discovering and adapting to the structural features in the input pattern.

In Reinforced learning method a supervisor is present but expected output is not present to the network. Its only indicates that either output is correct or incorrect. In this method a reward is given for correct output and penalty for wrong output.

Among all these methods Supervised and Unsupervised methods are most popular methods for training the network.

2.4 Activation Function

Activation function f is performing a mathematical operation on the signal output. Activation functions are chosen depending upon the type of problem to be solved by the network. There are some common activation functions[5] :

- 1) Linear activation function
- 2) Piecewise Linear activation function
- 3) Tangent hyperbolic function
- 4) Sigmoidal function
- 5) Threshold function

Linear activation function will produce positive number over the range of all real number [9].

Threshold function is either binary type or bi-polar type. If the threshold function is binary type its range is in between 0 to 1 and if the threshold function is bi-polar type its range lies between -1 to 1 [5].

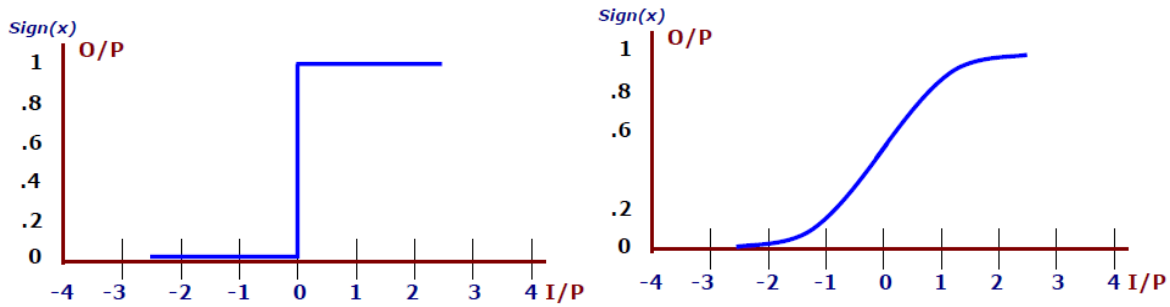


Fig.5 Threshold function

Piecewise linear function is also named as saturating linear function. This function has binary and bipolar range for saturating limits of the output. Its range lies between -1 to 1 [5].

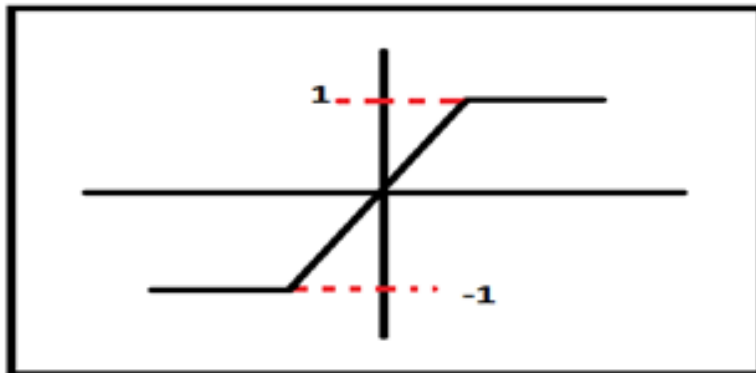


Fig.6 Piecewise linear function

Sigmoidal function is non-linear curved S-shaped function. This function is the most common type of activation function used to construct the Neural network. Sigmoidal function is strictly increasing function by nature [5].

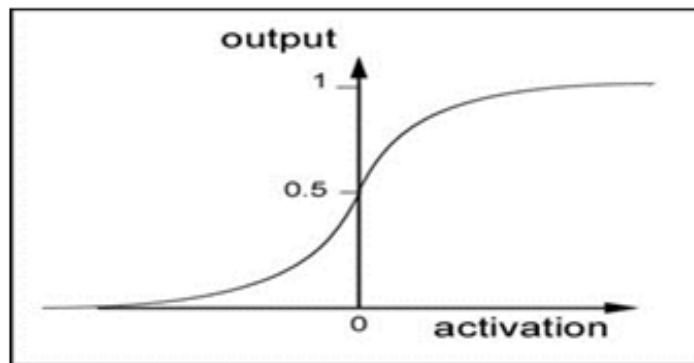


Fig.7 Sigmoidal function

3. RELATED WORK

3.1 *Hidden nodes and its effect*

As per describe earlier “Hidden nodes are nodes which are neither in the input layer nor output layer”. There are two types of effect occurred due to the hidden nodes [13].

Overfitting:

If there are so many neurons in the hidden layers it might cause Overfitting. Overfitting occurs when unnecessary more neurons are present in the network.

Underfitting:

If the number of neurons are less as compared to the complexity of the problem data it takes towards the Underfitting. It occurs when there are few neurons in the hidden layers to detect the signal in complicated data set.

3.2 *Hidden nodes selection in hidden layer*

Deciding the number of neurons in the hidden layer is a very important part of deciding your overall neural network architecture. Hidden layers do not directly interact with the external environment but still they have a tremendous influence on the final output. There are some various approaches to find out number of hidden nodes in hidden layer.

1. Try and Error Method[2]

Try and error method characterised by repeated, varied attempts which are continued until success or until the agent stops trying. This method divides in to two approaches [10].

Forward Approach : This approach begins by selecting a small number of hidden neurons. We usually begin with two hidden neurons. After that train and test the neural network. Then increased the number of hidden neurons. Repeat the above procedure until training and testing improved.

Backward Approach : This approach is opposite of Forward approach. In this approach we start with large number of hidden neurons. Then train and test the NN. After that gradually decrease the number of hidden neurons and again train and test the NN. Repeat the above process until training and testing improved.

2. Rule of thumb method[3]

Rule of thumb method is for determining the correct number of neurons to use in the hidden layers, such as the following:

- The number of hidden neurons should be in the range between the size of the input layer and the size of the output layer

- The number of hidden neurons should be $2/3$ of the input layer size, plus the size of the output layer
- The number of hidden neurons should be less than twice the input layer size

3. **Simple Method[3]**

It is a simple method to find out neural network hidden nodes. Assume a back propagation NN configuration is l-m-n. Here l is input nodes, m is hidden nodes and n is output nodes. If we have two input and two output in our problem then we can take same number of hidden nodes [3]. So our configuration becomes 2-2-2 where 2 is input nodes, 2 is hidden nodes, 2 is output nodes.

4. **Two phase method[3]**

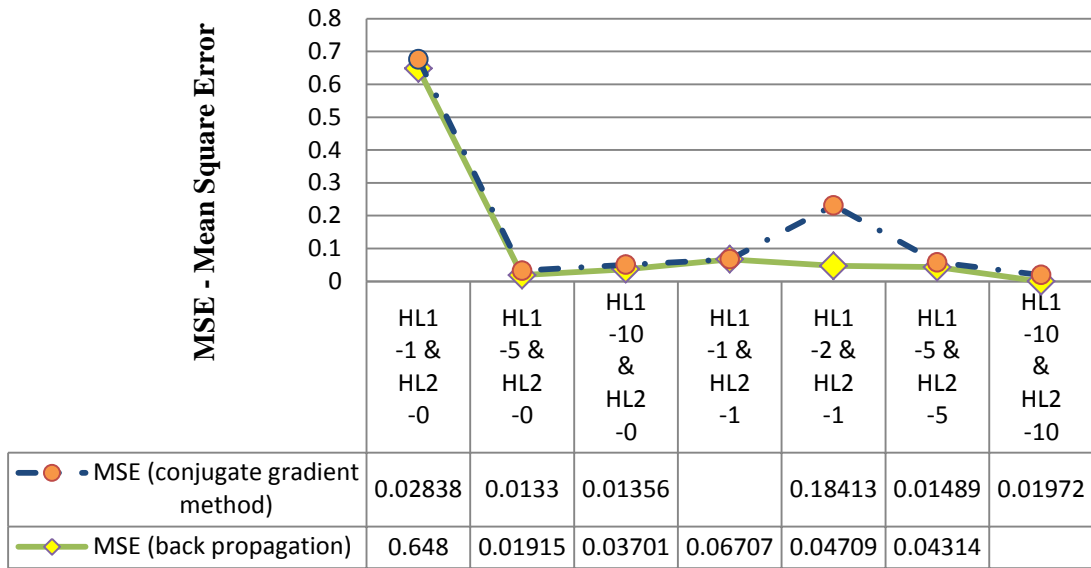
In two phase method the termination condition is same as the trial and error method but in a new approach. In this method data set is dividing into four groups. Among all four groups two groups of data are used in first phase to train the network and one group of remaining data set is used in second phase to test the network. Last group of data set is used to predict the output values of the train network. This experiment is repeated for different number of neurons to get minimum number of error terms for selecting the number of neurons in the hidden layer [3].

5. **Sequential orthogonal approach[7]**

Another approach to fix hidden neuron is the sequential orthogonal approach. This approach is about adding hidden neurons one by one. Initially, increase N_h sequentially until error is sufficiently small. When adding a neuron, the new information introduced by this neuron is caused by that part of its output vector which is orthogonal to the space spanned by the output vectors of previously added hidden neurons. An additional advantage of this method is that it can be used to build and train neural networks with mixed types of hidden neurons and thus to develop hybrid models

4. **EXPERIMENTS AND RESULTS**

We take one data set named Sales Forecasting and try to find out MSE with two different methods. First one is Back propagation method and second one is Conjugate gradient method. For that we start with single hidden layer with two neurons and then we increase the no.of neurons and train the network also note down the result. After that we add one another hidden layer and then we increase the no.of hidden nodes in each layer and train the network also note down the result.



No.s of Hidden nodes in Hidden layer

Fig.8 Experimental Graph

5. CONCLUSION

From above experiments and result we conclude that in Back propagation method from particular point the MSE error become steady but in conjugate gradient method the MSE is more fluctuate then back propagation. Also if suitable numbers of hidden nodes are taken we get better result in less training time. As we increase the number of hidden layers then accuracy can be obtained up to great extent but neural network became complex then previous. In this paper, methods for selecting number of hidden nodes found from literature are studied. Experiments are performed on real training data by changing the number of hidden nodes and layers. The obtained Mean Square Error for different experiments has been noted down and compared.

REFERENCES

- [1] S. N. Sivanandam, S. Sumathi, and S. N. Deepa, *Introduction to Neural Networks Usin Matlab 6.0*, Tata McGraw Hill, 1st edition, 2008.
- [2] “Review on Methods to Fix Number of Hidden Neurons in Neural Networks K. Ghana Sheila and S. N. Deepak, “Review on Methods to Fix Number of Hidden Neurons in Neural networks,” *Mathematical Problems in Engineering*, vol. 2013, Article ID 425740, 11 pages, 2013. doi:10.1155/2013/425740
- [3] Saurabh Karsoliya, “Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture,” *International Journal of Engineering Trends and Technology- Volume 3 Issue 6- 2012*
- [4] http://en.wikipedia.org/wiki/Types_of_artificial_neural_networks
- [5] http://www.myreaders.info/02_Fundamentals_of_Neural_Network.pdf
- [6] Boger, Z., and Guterman, H., 1997, “Knowledge extraction from artificial neural network models ,” *IEEE Systems, Man, and Cybernetics Conference*, Orlando, FL, USA

- [7] Berry, M.J.A, and Linoff, G. 1997, Data Mining Techniques, NY: John Wiley & Sons
- [8] Blum, A., 1992, Neural Networks in C++,NY:Wiley
- [9] P.SIBI, S.ALLWYN JONES, P.SIDDARTH,” Analysis of different activation functions using back propagation neural networks.”
- [10] <http://www.heatonresearch.com/articles/5/page2.html>
- [11] <http://www.eee.metu.edu.tr/~halici/courses/543LectureNotes/lecturenotes-pdf/ch9.pdf>
- [12] <http://www.cs.bham.ac.uk/~jxb/INC/112.pdf>
- [13] “Behaviour Analysis of Multilayer Perceptrons with Multiple Hidden Neurons and Hidden Layers,” International Journal of Computer Theory and Engineering, Vol. 3, No. 2, April 2011, ISSN: 1793-8201