Determining Hidden Neurons with Variant Experiments in Multilayer Perception using Machine Learning Neural Networks

K.Meenakshi Sundaram, S.Karthigai

Abstract: ***Neural network has broadly been employed in various fields for its efficacy and its superiority. Excellence results provided can be directly provided in various analyses. Besides the variant types of neural network, Multi layer perceptron plays a vital role for its adaptive learning ability. The network makes prediction based on learn of training set. Neural has three layers then the layers are the Input, Hidden and the Output Layers. There may be more than one hidden layer but there is one input and output layer. The hidden or the intermediate layer is considered as an engine of the complete network as it has the non linear activation function and they has a sensational domination in the finishing result . The amount of neurons in three layers determines the excellence of the network. The neuron in the input and the output layer is fixed as per the dataset while for the intermediate layer it is fixed by the user in random. Increase in neuron cause over-fitting while decrease cause under fitting and these assumptions have a great impact in the final outcome. This paper discusses the existing approaches for fixing the hidden neurons and proposes a method to fix the neurons in the intermediate layer and analyse the quality of the group. The proposed procedure has variant approaches to determine the hidden neuron and they are compared. The experiment is done in WEKA and the accuracy is checked with measures.***

Keywords: Data Mining, Multi Layer Perceptron, Hidden Neurons, WEKA.

I. INTRODUCTION

Data mining is the process [4] of revealing the concealed patters in the set. It is a process for assembling and managing info from numerous sources for the purpose of gaining in a single and detailed view. It has many procedures in separation to categorize, making groups or to find relation among the samples. The traditional procedures deals with the small dataset but now as the samples grew the traditional algorithm gives less accuracy so it has to be enhanced without affecting the quality of the result.

Data Mining Process

Data mining process has following includes following steps

A. Analysis

The decision makers need to frame the goals that the process expected to achieve. The problem and the objectives must be clearly stated. One cannot use procedures without the perception of the outcome.

B. Collection

This includes finding the best source for the data that is required. If the data is not available in the data warehouse, the user need to be identifying the required information, extract it and stored in some temporary system.

C. Preparation

Generally while creating the store, the data is accumulated from variant sources. So there is a possibility to encounter problems like missing value, conflicts and ambiguity. Also the same data set might be differed with variables. To avoid this circumstances data should be cleaned before processing.

D. Validation

The step takes the samples and applies a number of relevant techniques. For each one the results should be evaluated and interpreted. This is an iterative process which would lead to select one or more techniques that suits for further exploration and validation.

E. Visualization

Explaining the results to the decision makers is a vital procedure during the process. Most tools include data visualization modules. These tools communicate the results with more than two dimensions in a understandable format.

Neural Network

A processing system evaluated resembles the models of genetic neural [8] network. It is an adaptive structure that may change its arrangement by updating the weights that flows during the training time. It has "brain like" computer system with many interconnected processing units. There are variant networks architecture among them Multi layer perceptron has a significant role as it has more layers.

Multi Layer Perceptron

It belong to a feed forward artificial network consists of three layers of nodes. Other than the input and output a node, every node is a neuron that uses a non-linear activation function. It involves a supervised learning method known as ‘Back propagation’ for training. It is distinguished from the linear one by its multiple layers and non-linear activation. It can differentiate the data that is not linearly separable. The error and loss is estimated at the output and dispersed back via the intermediate layers. The most common used method is “Gradient descent” algorithm to adjust the weight.

Revised Manuscript Received on July 10, 2019.

Dr.K.Meenakshi Sundaram, Associate Professor of Computer Science, Erode Arts and Science College Erode -India.

S.Karthigai, Research Scholars in Computer Science, Erode Arts And Science College Erode – India.

Hidden layer

A middle [2] layer is between the put in and the productivity layer. The actual processing is done via weighted connections in this layer. Then link the processed information to the output layer. So, the number of hidden neurons that should be kept need to be carefully selected. The generalization of an ANN largely depends on the hidden layer. Too many hidden neurons over-train the network whist too few leads to insufficient learning. The data is linearly separable than there is no need to use hidden layer. But in case which deals with arbitrary decision of complicated densities, then more hidden units are needed.

The number of hidden units depends on the number of:

- Input and output units.
- Training cases.
- Noise in the targets.
- Complexity of the function to be learned.

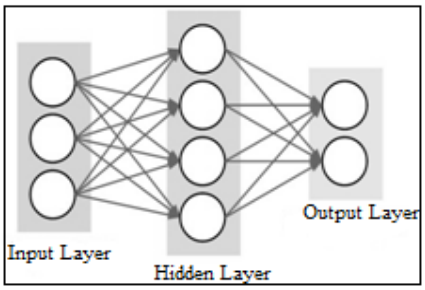

Figure 1. MLP Architecture

The Figure 1 shows the three layers Input, hidden and Output layer. This is hidden layer architecture with four nodes. The number of node in the hidden layer is in between the range of the input and output layers.

Lung Carcinoma

Lung Cancer are also called as Carcinoma is a cruel tumor identified by uninhibited [10] increase of tissue in the lung. The growth can enter by 'metastasis' into close by tissue or some other parts.

The types of Lung Cancer are Small-Cell lung carcinoma as SCLC and Non-Small-Cell lung carcinoma as NSCLC.

Most common symptoms for both types are

a. Coughing (including coughing up blood),
b. Tobacco smoking ,
c. Loss in weight,
d. Shortness in breath and
e. Pain in chest.

About ten to fifteen percent of belongings occur in the public who have never smoking habits. These cases are often caused by a combination of genetic factors, second-hand smoke and air pollution. The diagnosis of cancer is confirmed by biopsy. NSCLC needs surgery, whereas SCLC usually may be cured in chemo and radiotherapy.

Types of Lung Carcinoma

Two types of Lung Carcinoma can be explained are as follows

A. SCLC

The type of the malignant that has the most commonly resides within the lung, and occasionally arises in other sites. It has a lesser increasing time but higher growth rate, and earlier development of metastases. They usually presents in the central airways.

B. NSCLC

Non-SCLC have eighty five percent of all lung cancers. The most common types are squamous cell carcinoma, large cell carcinoma, and adeno carcinoma.

Adeno carcinoma are about forty percent which usually finds in the peripheral lung tissue.

Squamous-cell carcinoma causes about thirty percent. They occur close to large airways.

Large-cell carcinoma is about bout nine percent. The name implies the cancer cells are large, with increase in cytoplasm, large and conspicuous nucleoli.

Problem definition

The intermediate layer is considered as an engine of the whole network as it has the non linear activation function and they has a sensational domination in the final result. The number of hidden layer as well as the number of neurons in the hidden layer stabilizes the aspect of the outcome. There is limited number of accepted and qualified definitions for fixing the number of neurons. So there is a need to give a standard definition for the hidden layer neurons.

Objectives

- To eliminate the over fitting and under fitting issue.
- To give a standard definition for the hidden layer neurons.
- To lessen the complexity of the network.

The remaining section includes the following Section II describes the related works, Section III elaborates the methodology, Section IV shows the experimental results and listing away the performance of the enhancement, and Section V tells about the review and recruits the upcoming work.

II. LITERATURE REVIEW

A Multiple Hidden Layers extreme Learning Machine Method was introduced by Dong Xiao *et al* [1]. Extreme Method proposes a multi hidden layers which are obtaining the characteristics of parameters from the first hidden layer. l hidden neurons with the activation function (x) are in the three hidden layer network structure. By introducing a new method the parameters of the remaining hidden layers are obtained. Like the expected hidden layer output, learning machine method makes the actual output zero error approach. Many experiments on deterioration and classification are done based on this method. This results shows the proposed one achieve the satisfactory results as compared with two and some other multilayer.

Foram S. Panchal *et al* [3] propose a hidden layer node selection method. Sales Forecasting data set is taken and the analysis is done with the metric Mean Squared Error with two different methods. Back propagation is first one and Conjugate gradient method is second one. Initially the work starts with one intermediate layer and two neurons and then the total number of neurons are increased to train the

network. From the study it is shown, Back propagation method is sound but in gradient conjugate method the MSE is more vary. The analysis found that if suitable total numbers of intermediate nodes are taken the improved result is get with fewer training time. But if the total number of hidden layers is increased then exactness can be got to great level but network became more complex. The obtained MSE has been noted down and compared. Finally it is concluded that the total number of hidden nodes based on similarity among input data.

Gnana Sheela *et al* [7] designed a latest method to fix the intermediate neurons in networks for twist speed calculation. The survey made to fix the neurons in the intermediate layer in the neural networks. The random number of selection in the intermediate neurons might cause either more than or beneath fitting trouble. This paper elevates these problems using convergence theorem. To predict the intermediate neurons, 101 different criteria are experienced on the basis of statistical error. This outcome improves the exactness with slightest error. To verify the hidden model, simulation were conduct on real time twist data.

Mary Anne Rao *et al* [5] study the connection among the total number of intermediate layer neurons and the exactness of the classifier. The dataset used is the Breast Cancer Wisconsin Dataset. The input and target samples were randomly divided as seventy percent for training and fifteen for validation and balance for testing. Rule of thumb method is suggested in all network choices. Different hidden layer structures with number of the total neurons in analysed. From the result it is shown while increasing the hidden neurons number improves the ability of the network.

Wilson Castro *et al* [9] design a optimizing multilayer perceptron networks by evaluating three parameters. They are number of intermediate layers , neurons, and activation function with the metric of sum of squares error . The sample is taken from physicochemical and microbiological milk samples. The networks have 3 neurons in the input layer, 6 neurons in the output layer, three to 27 of neurons in the intermediate layer. The number of designs was determined using three factorial-type. Using MATLAB in 2015 has a logical sequence been calculated and implement using parallel computing techniques. The results of MLP show that intermediate layer and its number of relevant effect on error.

III. METHODOLOGY

Existing methodology

Four existing methods [2] are discussed elaborately are as follows

Method 1: Try and Error Method

The method repeatedly does varied attempts and is continued till satisfactory results appear. Two variant approaches of the try and error method consists are as follows

Forward

Forward begins by choosing a few numbers of neurons usually with two approaches. After the training and testing in one cycle the network the number of neuron is increased. This is repeated until result is improved.

Backward

This approach is just opposite to Forward Approach where great number of neurons is taken initially. After the training and testing the network the numbers of neuron is decreased and again repeat the above process until result is improved.

Method 2: Thumb Rule method

Find out the ideal number of neurons in the hidden layers. The rules are as follows for the number of hidden neurons should be

- The size of the input layer and the output layer of the range can be calculated.
- Two-third of the total of the input and the output layer should be maintained.
- The input layer is twice more than the Thumb rule.

Method 3: Method of Simple Approach

This approach is the simplest method to renovate the nodes. The simple method has designed as Input as L = Hidden as M = Output as N nodes. In some formation the input and the output has equal number of neurons in that case same number of hidden nodes is also apply.

Method 4: Orthogonal Sequential Approach

Orthogonal approach is about to raise the neurons by one. Add to the node sequentially until error is reduced. The new information introduced by this new node is caused when a node is added by the part of its output that is orthogonal to the space. Sequential method is used to construct and guide networks with varied types of neurons and hence aid in developing hybrid models.

Advantage of existing method

In every method accuracy will be high as it is repeated till the expected result is get.

Disadvantage of existing method

All the method is time consuming as the hidden neuron fixation is continued till the satisfactory result is reached.

Proposed Methodology as HMLP

The Proposed method HMLP are referred as Hidden-Multi Layer Perceptron in the expansion of Hidden Neuron suggests four approaches based on the number of attributes and classes.

The attributes are the variables used to define the samples and the class are the outcome of the sample that is defined on the basis of the relation between the variable and samples.

Approach – 1

The total number of neurons equal to the total number of attributes. If there are eight attributes then number of hidden neuron is eight.

$$\text{Hidden Neuron} = \text{Number of Attribute} \quad (1)$$

Approach – 2

The total number of neurons equal to the total number of classes. If there are three classes then number of neuron is three.

$$\text{Hidden Neuron} = \text{Number of Class} \quad (2)$$

Approach – 3

The total number of neurons equal to the sum of number of attributes and classes. If there are eight attributes and three classes then number of hidden neuron are eleven.

$$\text{Hidden Neuron} = \text{Number of (Attribute + Class)} \quad (3)$$

Approach – 4

The total number of neurons equal to the half of the sum of number of attributes and classes. If there are eight attributes and three classes then number of neuron is five.

$$\text{Hidden Neuron} = \text{Number of (Attribute + Class)} / 2 \quad (4)$$

Procedure HMLP

Step 1: Initialize the neural network with m inputs, n outputs and e connections.
Step 2: Split the data set into N Folds.
Step 3: Assign N-1 folds for training and the rest for testing.
Step 4: Begin first Fold.
Step 5: Initialize the input layer and connect it to the hidden layer as on (1 - 4) with assigned weights randomly.
Step 6: Propagate forward to generate the output.
Step 7: Calculation of the cost term over n training samples by,

$$E = \frac{1}{2n} \Sigma \| (y(x) - y'(x)) \|^2 \quad \text{--- (5)}$$

Step 8: Propagate the output activations back through the network in order to generate the deltas.
Step 9: The input activation output and delta are multiplied to find the gradient to find the new weight.
Step 10: The weight is updated to reduce the error.
Step 11: Repeat for N Folds.

Advantages

- The proposed method defines the hidden neuron in a well explained manner.
- The method gives four approaches and one of this must definitely suits to any kind problem.
- The method reduce the complexity by lessen the time while analysing the number of neuron.

IV. EXPERIMENTAL RESULTS

The database is created in Microsoft excel sheet. It is given by a medical practitioner. The dataset is enhanced with iterative optimization method [10] and then the hidden layer is expanded by the proposed procedure of this paper. The results are validated in WEKA tool. Nine evaluation metrics are used.

A. Data set

The Lung cancer dataset are collected from a medical practitioner. It consists of 15 attributes with 3772 instances.

Attributes

The fifteen attributes are Patient id, gender, chronic cough, Hemoptysis, Pain in chest, Dysponia, Cachexia, Infection in lungs, Swelling, Wheezing, Dypsnea, Clubbing in nails, Dysphasia, Tumor location and a class label with four classes Adeno carcinoma, Squemous carcinoma, Large cell Carcinoma and Small cell Lung Carcinoma..

B. PreProcessing

Pre processing is an earlier stage [12] in mining the data to clean, integrate, select and reduction of the set. In this work Gain Ratio attribute evaluation pre processing method is carried out and ten attributes are selected based on the information gain. The selected attributes are Patient id, gender, Hemoptysis, Dysponia, Cachexia, Wheezing, Dypsnea, Dysphasia, Tumor location and class label with four trials Tumor location as the out the lungs, anywhere in the lungs, in the bronchi, and in the centre of the lungs. So the input attribute are twelve.

C. Results of HMLP

Table 1 Summary

Approach	No. of hidden neuron	RMSE	Time in sec.
1	12	0.1028	85.78
2	4	0.1068	24.60
3	16	0.1080	103.22
4	8	0.1021	55.14

The Table 1 Shown above summary of the four proposed approaches. The quantity of unseen neuron is calculated as by the equation 1 to 4.

Among the four, Approach 4 (4) gives less RMSE. Though the Approach 2 (2) has less processing time it is not considered as good as it has high RMSE than approach 4. The lowest RMSE will give the highest accuracy.

Hence the fourth approach (4) is selected and is evaluated in the multi layer perceptron for further analysis.

Table 2. Comparision of HMLP with existing approach

Evaluation Measures	MLP	HMLP
TP Rate	0.948	0.973
FP Rate	0.024	0.012
Precision	0.950	0.974
Recall	0.948	0.973
F-Measure	0.947	0.973
MCC	0.931	0.962
ROC	0.996	0.997
PRC	0.992	0.994
Accuracy	94.8%	97.2%

Table 2 list the performance of proposed approach 4 in MLP with the existing approach.

D. Performance Analysis

Chart 1

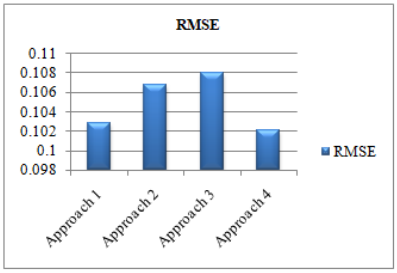

Chart 1 RMSE

The above Chart 1 shows that Approach 4 has the least Root Mean square error.

Chart 2

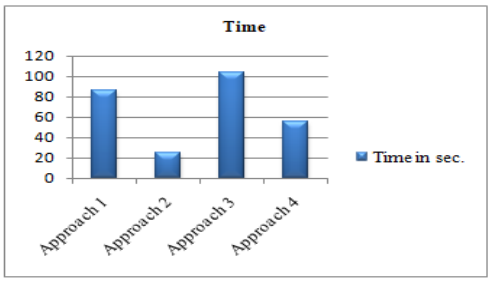

Chart 2 Processing Time

The above Chart 2 shows Approach 2 has minimum processing time as the number of neuron is equal to the number of classes, it has only four neurons. Approach 4 has moderate time as compared with the Approaches 2 and 3.

Chart 3

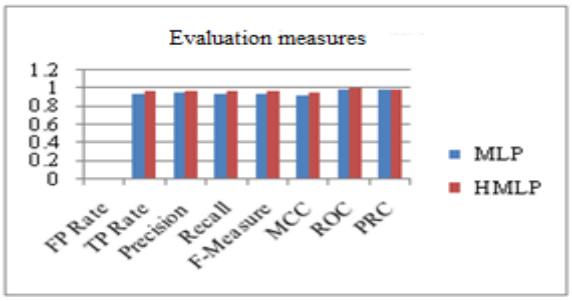

Chart 3 Evaluation measures

The above Chart 3 shows the Approach 4 with MLP perform better.

Chart 4:

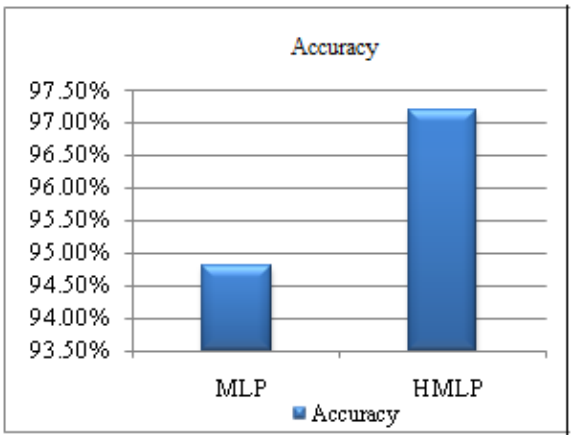

Chart 4 Accuracy

The above Chart 4 shows the accuracy and is concluded the proposed approach has high accuracy.

V. CONCLUSION AND FUTURE WORK

The Multi Layer Perceptron has the output, the intermediate are called as hidden and the input layers are of three types. The total number of intermediate layer and intermediate neurons are fixed through the user and they have a great impact in the result. This paper proposes a new method HMLP with one hidden layer and four approaches to determine the number of neurons. Lung cancer data set is taken for the analysis. The evaluation metrics are used for comparison. From the results it is shown Approach 4 has the least RMSE with highest accuracy.

In future, this hidden layer implementation is used with variant dataset for analysis and can implemented for CNN as convolutional neural network

REFERENCES

1. Dong Xiao, Beijing Li, and Yachun Mao, "A Multiple Hidden Layers Extreme Learning Machine Method and Its Application", Hindawi Mathematical Problems in Engineering Volume 2017.
2. Blum, A., 1992, Neural Networks in C++, NY:Wiley.
3. Foram S. Panchal , Mahesh Panchal, "Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural Network", IJCSMC, Vol. 3, Issue. 11, November 2014.
4. Berry, M.J.A, and Linoff, G. 1997, Data Minning Techniques, NY: John Wiley & Sons.
5. M. Gethsiyal Augasta, T. Kathirvalava kumar. Pruning algorithms of neural networks - a comparative Study Central European Journal of Computer Science, Computer Applications Department, Sarah Tucker College, Tirunelveli, 627007, TN, India, Department of Computer Science, V.H.N.S.N College, Virudu Nagar, 626001, TN, India 06 September 2013.
6. P.Sibi, S.Allwyn Jones,P.Siddarth,"Analysis of different activation functions using back propagation neural networks."
7. K. Gnana Sheela and S. N. Deepa, "Review on Methods to Fix Number of Hidden Neurons in Neural Networks", Hindawi Publishing Corporation Mathematical Problems in Engineering Volume 2013.
8. Mary anne roa, Proceso fernandez jr., "Study on optimization of artificial neural network generalization power based on architecture", International Journal of Soft Computing and Artificial Intelligence, Volume-5, Issue-1, May-2017.
9. Wilson Castro, Jimy Oblitas, "Multilayer perceptron architecture optimization using parallel computing techniques", Plos one, V12 (12), 2017.
10. S.Karthigai, Dr.K.Meenakshi Sundaram "Iterative Supervised Learning Method For The Categorization Of Lung Carcinoma", International Journal of Research and Analytical Reviews, Volume 5, Issue 3, September 2018.
11. Karthigai S , Meenakshi Sundaram K, " Performance of Lung Carcinoma in Classification Neural Network with Pre Processing using WEGA", International Journal of Future Revolution in Computer Science and Communication.
12. www. Lungcancer.wikipedia.

AUTHORS PROFILE

Dr.K.Meenakshi Sundaram having 31years of teaching experience. He produced 45 M.phil research Scholars in computer science. Also 6 Ph.D Public V:Vavoce Examination as examiner. He scored as a member in Board of Studies of Various universities and Autonomous Colleges. He Published 45 papers in International Journals and presented 60 papers in National / International conferences. His Area of Interest is Data mining and software Engineering.

S.Karthigai is currently pursuing Ph.D as a part time research scholar under the guidance of Dr.K.Meenakshi Sundaram in the Department of Computer Science, Erode Arts and Science College [Autonomous], Erode – Tamilnadu - India. She has obtained her Masters degree in Computer Applications from Shrimathi Indra Gandhi College, Trichy under Bharathidasan University and M.Phil degree in Computer Science at Bharathidasan University - Trichy. Her area of interests includes, Data Mining and Artificial Intelligence. Shee has presented Two papers in National and five International level conference. She published five papers in International Journals. She published "Linux and Shell Programming" at Charulatha Publications