International Journal on Emerging Technologies (Special Issue NCETST-2017) **8**(1): 660-666(2017)
(Published by Research Trend, Website: www.researchtrend.net*)*

ISSN No. (Print) : 0975-8364
ISSN No. (Online) : 2249-3255

Comparison Analysis of single Multiplicative neuron with Conventional Neuron Models

Pankaj Kumar Kandpal and Ashish Mehta

Department of Computer Science, Kumaun University, Nainital, (U.K.), India

ABSTRACT: In this paper, Multiplicative Neuron Models is used for classification of nonlinear problems. The conventional neuron model, "Multi Layer Perceptrons (MLP) is been taken for the comparative analysis with Multiplicative Neuron Model. It is found that Multiplicative neuron model with single neuron, is sufficient for classification that require number of neurons in different hidden layers of a conventional neuron network. For comparative analysis of both models, various parameters of Artificial Neural Network like learning rate, execution time, number of iteration, time elapse in training, mean square error etc. are considered. After comparing the various above mentioned parameters, it is observed that execution time, number of iteration, time elapse in training is minimum in the case of Multiplicative neuron model. On the basis of results over two datasets IRIS and Mammografic mass, it is observed that Multiplicative neuron model performance is better for classification. Our study also justifies the earlier studies done by Deepak Mishra and *et. al*. [12][13].

Index: Multiplicative Neuron, Multilayer Perceptron, Classification

I. INTRODUCTION

Artificial Intelligence is the branch of the computer science concerned with the study and creation of computer systems that exhibit some form of intelligence: system learn new concepts and tasks, system that can reason and draw useful conclusion about the world around us, system that can understand a natural language or perceive and comprehend a visual sense, and system that perform other types of feats that require human types of intelligence [1]. The Artificial Neural Networks is one stream of Artificial Intelligence.

Artificial Neural Networks is the mathematical model of biological neurons. Although all these models were primarily inspired from biological neuron, after giving the so many contribution by plenty of researchers still a gap between philosophies used in neuron models for neuroscience studies and those used for artificial neural networks (ANN). Some of neural network models exhibit a close correspondence with their biological counterparts whiles other far away with their counterparts. It is being contributed by several scientists that gap between biology and mathematics can be minimized by investigating the learning capabilities of biological neuron models for use in the applications of classification, time-series prediction, function approximation, etc. In this paper, compared the two very efficient models and after analyzing the results, it is found that which one is the better model in context of various parameters of Artificial Neural Network like Learning Rate, Execution Time, Number of Iterations, Time Elapse in training etc.

The first artificial neuron model was proposed by McCulloch and Pitts [7] in 1943. They developed this neuron model based on the fact that the output of neuron is 1 if the weighted sum of its inputs is greater than a threshold value, and 0, otherwise. In 1949, Hebb[8] proposed a learning rule that became initiative for ANNs. He postulated that the brain learns by changing its connectivity patterns. Widrow and Hoff [9] in 1960 presented the most analyzed and most applied learning rule known as least mean square rule. Later in 1985,Widrow and Sterns [10] found that this rule converges in the mean square to the solution that corresponds to least mean square output error if all input patterns are of same length. A single neuron of the above and many other neuron types proposed by several scientists and researchers are capable of linear classification [11]. In this paper in, the II part we have exhibited biological neuron.

II. BIOLOGICAL NEURON MODEL

A. Multilayer Perceptron

It is a very well known conventional model. The adapted perceptrons are arranged in layers and so the model is termed as multilayer perceptron. This model has three layers: an input layer, an output layer, and a layer in between, not connected directly to the input or output, and hence called the hidden layer.

For the perceptrons in the input layer, linear transfer function is used, and for the perceptrons in the hidden layer and the output layer, sigmoidal or squashed-S functions is used. The input layer serves to distribute the values they receive to the next layer and so does not perform a weighted sum or threshold. The input-output mapping of multilayer perceptron is shown in Fig. 1.

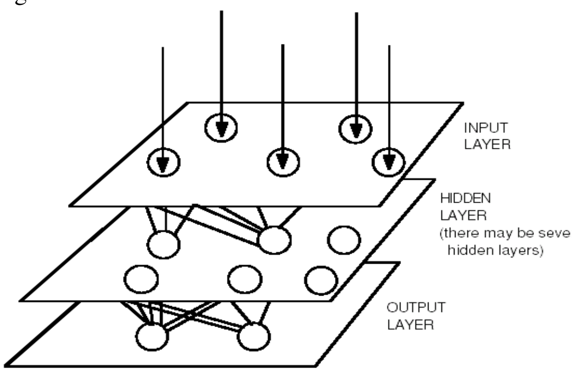

Fig. 1. Multilayer Network.

Many capabilities of neural networks, such as nonlinear functional approximation, learning, generalization etc. are, in fact, due to nonlinear activation function of each neuron. Sigmoid Activation Function is given below:

h1 = -neth1

The activity of neurons in the input layers represents the raw information fed into the network, the activity of neurons in the hidden layer is determined by the activities of the neuron in the input layer and connecting weights between input and hidden units. Similarly, the activity of the output units depends on the activity of neurons in the hidden layer and the weight between the hidden and output layers. This structure is interesting because neurons in the hidden layers are free to conduct their own representation of the input. [2]

B. Multiplicative Neuron Model

Only single neuron of this model is used for the classification task. In this model, aggregation function is based upon the multiplicative activities (Ω) at the dendrites, instead of summation activities given in the fig.1.

$$\mathbf{\Omega}(\mathbf{x}, \boldsymbol{\theta}) = \prod_{i=i}^{n} (w_i x_i + b_i)$$

In above given equation Ω is a multiplicative operator with weights w_i, x_i inputs and biases bi. In the given equation ∏ (production) is being used instead of ∑ summation. It is investigated the complexity of computing and learning for multiplicative neuron.

In particular, we derive upper and lower bounds on the Vapnik- Chervonenkis (VC) dimension and pseudo dimension for various types of networks with multiplicative units [20-22]. In the Internal architecture and computation methods are different but the procedure of training; testing and prediction are same as used in Multi-Layer Perceptron model. Unlike the higher-order neuron, this model is more simpler in terms of its parameters and one does not need to determine the monomial structures prior to training of the neuron model. Multiplicative Neuron Model is used for problems with high nonlinearity and it can be trained easily.

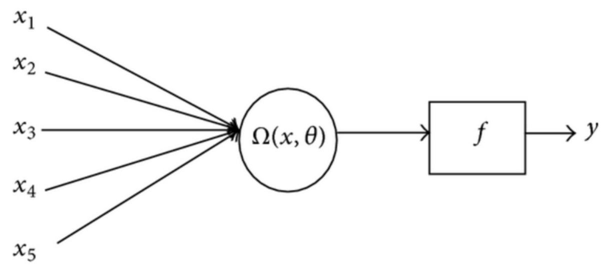

Fig. 2. A Generalized Single Neuron.

III. LEARNING FOR MULTIPLICATIVE NEURON MODEL

In the same way as multilayer layer perceptrons, multiplicative neuron model uses error back propagation learning. The simplicity of back propagation methods make it convenient for the models to be used in different situation, unlike the high order neuron model, which is difficult to train and is susceptible to combinatorial explosion of terms. A simple gradient descent rule, using a norm-squared error function, is described by the following set of equations.

$$\text{net} = (w_1 x_1 + b_1) * (w_2 \quad x_2 + b_2) \ldots.. \qquad (1)$$

$$e = \frac{1}{2}(y - t)^2 \qquad (2)$$

$$\frac{de}{dy} = (y - t) \qquad (3)$$

$$\frac{dy}{dnet} = y(1 - y) \qquad (4)$$

$$\frac{dnet}{dw_i} = \frac{net}{(w_i x_i + b_i)} x_i \qquad (5)$$

$$\frac{de}{dw_i} = \frac{de}{dy} \frac{dy}{dnet} \frac{dnet}{dw_i} \qquad (6)$$

$$\frac{de}{dw_i} = (y - t) * y * (1 - y) * \frac{net}{(w_i x_i + b_i)} \, x_i \quad (7)$$

$$\frac{de}{db_i} = \frac{de}{dy} \frac{dy}{dnet} \frac{dnet}{db_i} \quad (8)$$

$$\frac{dnet}{db_i} = \frac{net}{(w_i x_i + b_i)} \quad (9)$$

$$\frac{de}{db_i} = (y - t) * y * (1 - y) * (w_i x_i + b_i) \quad (10)$$

$$w_{i(new)} = w_{i(old)} - \frac{de}{dw_i} * eta \quad (11)$$

$$b_{i(new)} = b_{i(old)} - \frac{de}{db_i} * eta \quad (12)$$

In the first equation productive function being used as an activation function. In the Equ.2 deviation between actual value (y) and target value (t), where is eta(□) is learning rate which can be assigned a value on the heuristics basis. After using back propagation weight and bias gets new improved values after every epoch (equ.11-12). Using the back propagation learning method, it being solved some most popular classification problem in next section.

IV. COMPARATIVE ANALYSIS OF CLASSIFICATION

Two classification models have been selected and well known datasets has been taken. The experimental parameters show that multiplicative neuron and multi layer perceptron (MLP) has been trained by using two well known data sets IRIS and Mammografic Mass. Whole data set is been used for training and a small subset is been used for testing. The training is continued until the network going on improving. When network trained, the training is stopped. Training can be stopped in another condition when training goal in term of MSE is met or given iteration (epoch) are completed. . For simulation of problem the minimum configrational requirement of the computer is Pentium 4 processor with 2.3 GHz and 512 MB RAM.

A. IRIS Dataset

Iris data set is very popular dataset among researchers. It is open for all at university of California archive [17], having three species of Iris flower setosa, versicolor, virginica. Each flower has parts called petals & sepals, length and width of sepal & petal can be used to determine iris type. Data collected on large number of iris flowers. Neural net will be trained to determine specie of iris for given set of petal and sepal width and length. The authors compare the performance of multiplicative neural networks (MNM) with that of multi layer perceptrons (MLP). For this objective, the MLP taken with three layers, with multiple hidden neurons. The Fig.3 shows the mean square error (MSE) versus number of epochs (Iteration) curve for training with multiplicative neuron model (MNM) and MLP while dealing with the IRIS flower classification problem. It is cleared with the curve that Multiplicative neuron model with single neuron, learns easily and minimize the error early in comparison to multilayer perceptron. The table.2 exhibit the comparison between MLP and MNM in terms of deviation of actual outputs from corresponding targets. It can be seen with the help of table1 that the performance of MNM is better than that of MLP. From table1, it is observed that the training time required by MNM is much less than MLP. It means that a single neuron in MNM is capable to learn IRIS relationship almost four times faster than MLP with 18 hidden neurons. Table shows the comparison of training and testing performance with MLP and MNM, while solving the IRIS classification problem.

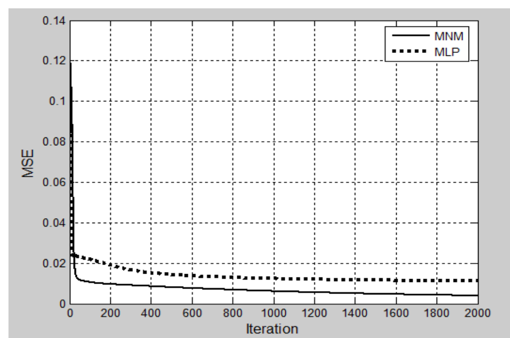

Fig. 3. Mean square error vs. iteration for training for IRIS problem.

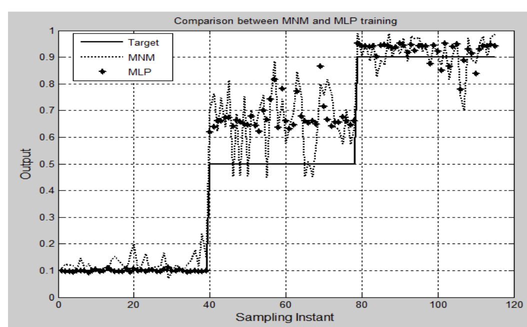

Fig. 4. Comparison between MNM and MLP training.

S.No.	Parameter	MNM	MLP
1	Training goel, in term of MSE (error check)	0.0001	0.00001
2	Iteration needed	500	4000
3	Training time in seconds	19	92
4	testing time in seconds	0	0
5	MSE for training data	0.0074	0.0058
6	MSE for testing data	0.0054	0.0033
7	RSME for training	0.858	0.0763
8	RMSE for testing	0.755	0.0572
9	Correlation coefficient	0.9682	0.9699
10	percentage of miss classification	5%	5%
11	number of neurons	1	23
12	learning late (□)	1.8	2.1

Table 1: Comparison of training and testing performance for IRIS problem.

Table 2, shows the input values and equivalent outputs values of both models. Figure 4 and figure 5 shows the training and testing results of IRIS datasets. The figures show that some marginal overlapping all three classes are clearly separable with each others.

Table 2: Comparison of Output of MNM and MLP for IRIS Problem.

Input	Target	Actual Output with MNM	Actual Output with MLP
0.678, 0.467, 0.656,0.833	0.9	0.8818	0.92848
0.544, 0.567, 0.724,0.867	0.9	0.95001	0.95041
0.167, 0.567, 0.154, 0.167	0.1	0.097149	0.092123
0.411, 0.7, 0.195, 0.167	0.1	0.15513	0.10947
0.256, 0.2, 0.412, 0.4	0.5	0.66848	0.64186
0.389, 0.267, 0.493, 0.433	0.5	0.56815	0.6241

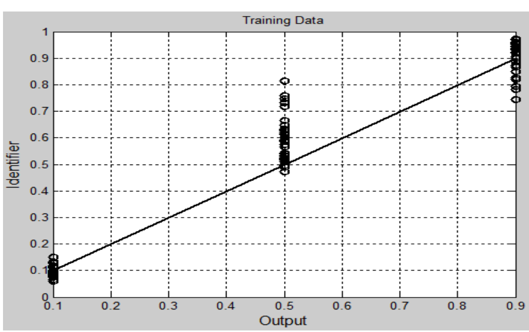

Fig. 5. Training results for IRIS Problem.

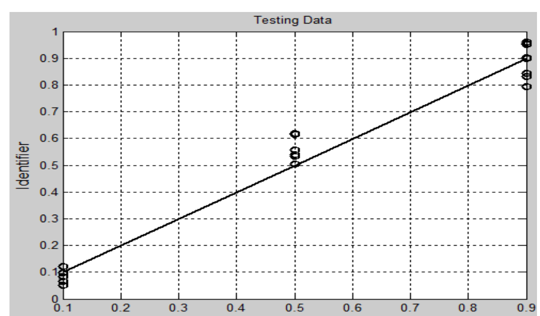

Fig. 6. Testing results for IRIS Problem.

B. Mammografic Mass Dataset

Mammography is the most effective method for breast cancer screening available today. However, the low positive predictive value of breast biopsy resulting from mammogram interpretation leads to approximately 70% unnecessary biopsies with benign outcomes. To reduce the high number of unnecessary breast biopsies, several computer-aided diagnosis (CAD) systems have been proposed in the last years.

These systems help physicians in their decision to perform a breast biopsy on a suspicious lesion seen in a mammogram or to perform a short term follow-up examination instead.

This data set can be used to predict the severity (benign or malignant) of a mammographic mass lesion from BI-RADS attributes and the patient's age [17]. The mammografic problem deal with the classification between benign (0) and malignant (1), authors compared the performance of multiplicative neural networks (MNM) with that of multi layer perceptrons (MLP). Depicted in the Fig. 7 that MSE versus epochs curves for training with MNM and MLP while dealing the problem. Where MLP takes 4000 epochs to learn the pattern, on the other hand, MNM takes only 1000 epochs. From the table 3, it is observed that the training time required by MLP is much more than MNM. It means that a single neuron of MNM is capable to learn mammographic mass pattern, where MLP model required 31 neurons.

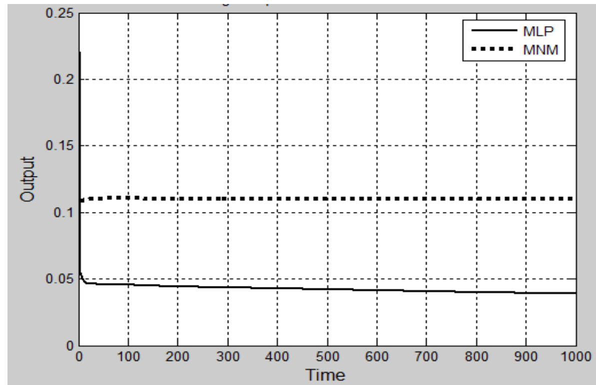

Fig. 7. Mean square error vs. iteration for training for Mammografic Mass problem.

Table 3: Comparison of training and testing performance for Mammografic Mass problem.

S.No.	Parameter	MNM	MLP
1	Training goel, in term of MSE (error check)	0.0001	0.00001
2	Iteration needed	1000	4000
3	Training time in seconds	106	214
4	testing time in seconds	0.18	0.02
5	MSE for training data	0.0606	0.0363
6	MSE for testing data	0.0663	0.0442
7	RSME for training	0.2462	0.1904
8	RMSE for testing	0.2575	0.2103
9	Correlation coefficient	0.5711	0.7509
10	percentage of miss classification	23%	13%
11	number of neurons	1	31
12	learning late (□)	0.77	0.85

Table 4, exhibits the comparison between MNM and MLP in terms of deviation of actual output from coresponding targets. In context of mammografic dataset results are not as better as IRIS problem but reveals a clear cut difference between the MLP and MNM.

Table 4: Comparison of Output of MNM and MLP for Mammografic Mass Problem.

Input	Target	Actual Output with MNM	Actual Output with MLP
0.172, 0.5,0.1,0.9,0.633	0.9	0.83568	0.94885
0.172,0.694,0.1,0.7,0.633	0.9	0.63788	0.87049
0.158,0.284,0.633, 0.1,0.633	0.1	0.14166	0.1139
0.158, 0.580, 0.366, 0.1, 0.366	0.1	0.15231	0.10555

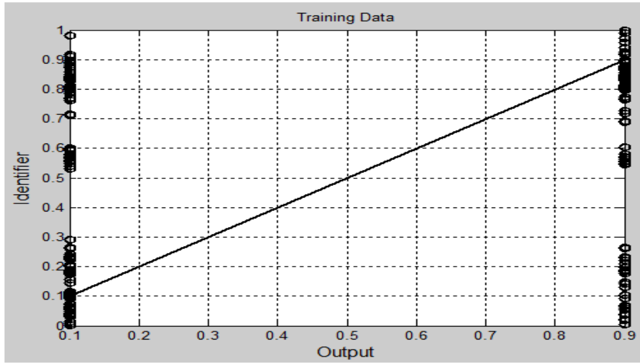

Fig. 8. Training results for Mammografic Mass Problem.

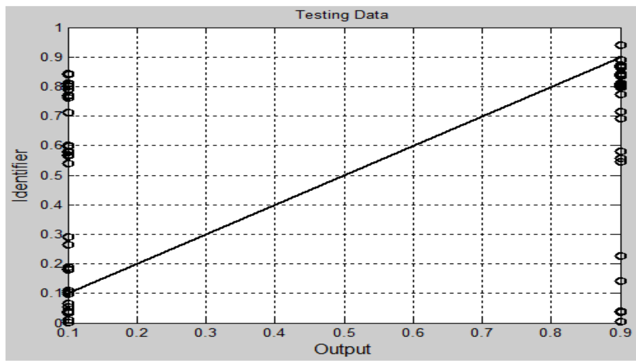

Fig. 9. Testing results for mammografic mass Problem.

V. CONCLUSION

After the finding the training and testing results of MNM and MLP using both popular IRIS and mammografic mass classification problem it can be percept that single multiplicative neuron capable of performing classification task as efficiently as a multilayer perceptron with many neurons and in the IRIS problem case its learning is even better than that of multilayer percedptron. It is also percept that training and testing time in case of MNM are significantly less as compared with MLP, in both problems. Therefore it is justified that MNM is better than MLP.

REFERENCES

[1] Dan W Patterson "Introduction to Artificial Intelligence and Expert System" Prentice Hall of India Private ltd 2005.

[2] Jacek M. Zurada "Artificial Neural System" West Publishing Company.

[3] Abhishek Yadav "Dynamical aspects and learning in Biological Neuron Models" department of electrical engineering IIT Kanpur June 2005.

[4] M. Balasubramanian, M. Fellows and V. Raman, unpublished

[5] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve", *Journal of Physiology*, **117**, pp. 500–544, 1952.

[6] W. J. Freeman, "Why neural networks dont yet fly: inquiry into the neurodynamics of biological intelligence", *IEEE International Conference on Neural Networks,* 24-27 July 1988, pp.1-7, vol.**2**, 1988.

[7] W. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics,* vol. **5**, pp. 115-133, 1943.

[8] D. Hebb, "Organization of behavior", John Weiley and Sons, New York, 1949.

[9] B. Widrow and M. E. Hoff, "Adaptive switching circuits", IREWESCON Connection Recors, IRS, New York, 1960.

[10] B. Widrow and S. Steams, "Adaptive signal processing", Prentice-Hall, Englewood Cliffs, NJ., 1985.

[11] M. Sinha, D.K. Chaturvedi and P.K. Kalra, "Development of flexible neural network", *Journal of IE(I),* vol.**83**, 2002.

[12] Deepak Mishra, Abhishek Yadav, & Prem K. Kalra, A Novel Neural Network Architecture Motivated by Integrate-And-Fire Neuron Model Department of Electrical Engineering Indian Institute of Technology Kanpur, India

[13] Deepak Mishra, Abhishek Yadav, & Prem K. Kalra, A Novel Multiplicative Neural Network Architecture Motivated by Spiking Neuron Model Department of Electrical Engineering Indian Institute of Technology Kanpur, India

[14] A. Yadav, D. Mishra, R.N. Yadav, S. Ray, P.K. Kalra, Time-series prediction with single integrate-and-fire neuron, *Science Direct, Applied Soft Computing* **7** (2007) 739–745.
[15] R N Yadav, P K Kalra, S John, Time series prediction with Single Multiplicative Neuron Model, *Science Direct, Applied Soft Computing* **7** (2007) 1157–1163
[16] Deepak Mishra, Abhishek Yadav, Sudipta Ray, Levenberg-Marquardt Learning Algorithm for Integrate-and-Fire Neuron Model, IIT Kanpur.
[17] https://archive.ics.uci.edu/ml/datasets.html
[18] Peter Dayan and L F Abbott, Theoretical Neuroscience "Computational and Mathematical modeling of Neural System, MIT Press Cambridge, London.
[19] Christof Koch, Biophysics of Computation " information processing in single neuron", Oxford University Press.
[20] N. sauer,on the density of family of sets, *J. Comb. theory*(A) **13** (1972) 145-147.
[21] D Haussler, P M Long, A generalization of Sauer's Lemma, *J. Comb. Theory*, **71** (1995) 219-240.
[22] P Auer, P M Long, W. Maass, G J Woeginger, On the complexcity of function learning, *Mac. Learn.* **18**(1995) 189-236.