IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-5, May 2018 E-ISSN: 2347-2693

Analytical Observation for classification of Multilayer Neuron Models using different datasets

Pankaj Kumar Kandpal[1*], Ashish Mehta[2]

[1*] Department of Computer Science / Kumaun University, Nainital,Uttrakhand, India
[2] Department of Computer Science / Kumaun University, Nainital,Uttrakhand, India

**Corresponding Author: kandpalzee@gmail.com , Tel.: +91-9412097979*

Available online at: www.ijcseonline.org

Accepted: 16/May/2018, Published: 31/May/2018

Abstract- In this paper, Multilayer Neuron model is used for classification of nonlinear problems. This conventional neuron model, is been taken for the analysis of while using different data sets. It is found, the Multilayer Neuron model showing its varying efficiency according to pattern of dataset. For analysis of model, various parameters of Artificial Neural Network like numbers of hidden neuron, number of attributes, learning rate, correlation coefficient, numbers of iteration, time elapse in training, mean square error etc. are being taken. After the analytical observation considering above various mentioned parameters, it is observed that there is no thump rule on behalf we can say that Multilayer Neuron Model follow the particular rule. The learning of model depends on the pattern of the dataset and the quality of data.

Keywords— Multilayer Neuron, Classification, , analysis, Class.

I. Introduction

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]. Multilayer neuron model is very popular and simple model. Various researches have been explored to solve the difficult classification and function approximation problems. In this proposed model we considered the analytical observation on behavior of Multilayer Neuron Model. There are some papers over classification problem presented by us using different model[22-24]. In this paper in the I section contains introduction

Of research work and paper study, the II section of paper we have exhibited Multilayer neuron model. III section describing different datasets used in the paper, IV section is results and discussion and in last and V section concludes the research works and future direction.

II. BIOLOGICAL NEURON MODEL

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 are 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 Figure1.

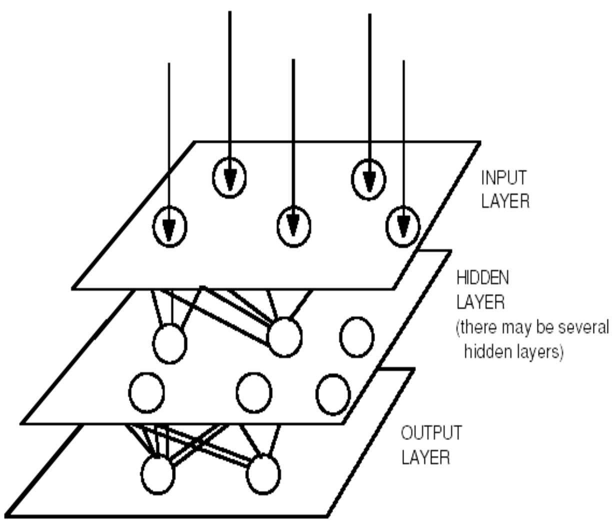

Figure 1. Multilayer Neural Network.

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:

$$h_1 = -net_{h1} \quad (1)$$

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]

III. DATASET USED

A. Iris Dataset

Iris data set is very popular dataset among researchers as Fisher Iris. It is open for all at university of California archive, 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.

B. Wine dataset

Wine dataset is using chemical analysis determine the origin of wine. It is generated by Forina, M. et al, PARVUS - An Extendible Package for Data Exploration, Classification and Correlation. Institute of Pharmaceutical and Food Analysis and Technologies, Via Brigata Salerno, 16147 Genoa, Italy.

These data are the results of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars. The analysis determined the quantities of 13 constituents found in each of the three types of wines. There are total 178 instants with no missing value.

C. Brest cancer data set

Total ten inputs, Brest cancer dataset donate by Dr. William H. Wolberg to UCI. This dataset is the study of that was conducted at the university of Wisconsin Hospital, Madison, about patient who had undergone surgery for Brest Cancer. This task is to determine if the tumor is benign otherwise malignant.

D. Mammographic Mass Data set

Mammography is the most effective method for Brest cancer screening available. However the low positive predictive value of Brest biopsy resulting from mammogram Dhankar Singh Verma, koganga Purkayastha approximity 70% unneccessry piopsies with benign outcomes. To reduce the high number of unnecessary Brest biopsies, several computer added diagnosis system have been produced in last few years. The system helps physicians in their decision to perform a Brest biopsy on the suspicious lesion seen in a mammogram or to perform a short term follow up examination instead. This dataset can be used to predict the severity of mammographic mass lesion from BI-RADS and patient age.

IV. RESULTS AND DISCUSSION

Analytical Observation for Classification

In this paper we tried to explore that multilayer Neuron model‘s adaptation and classification properties while using different types of datasets. In this paper the observation in respect to datasets categorized in two segments. One, the data are classified in two classes like Brest Cancer and Mammographic mass dataset. Second, the data are classified in three classes like wine and Iris dataset. All the four datasets are taken from UCI repository. The authors have taken the best results of all the datasets for study. The number of iterations has taken same in all classification problems. For each simulation the minimum requirement of hardware configuration is Pentium 4 processor with 1.8 GHz and 512 MB RAM.

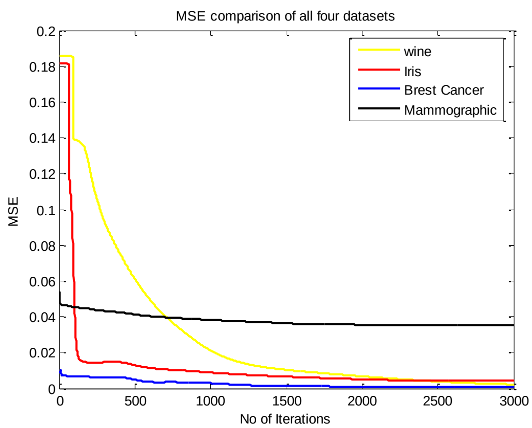

Figure2. Mean square error vs. iteration for training of all classification problems.

A. Two classes classification problems

The Brest cancer Wisconsin problem and Mammographic mass problem having two types of classes (benign otherwise malignant). In the graph depicted in the Fig. 2 that blue line of Brest cancer Wisconsin problem trained better way and classified properly than that of Mammographic mass problem. MSE (red line) for Mammographic mass problem lie away from the zero line where the MSE for Brest Cancer problem (blue line) near to zero line, it means that Brest cancer problem error minimize the error properly. The learning of Mammographic mass problem is poorer than learning of Brest cancer. Depicted in the table1, classification percentage of Brest Cancer is 100% where classification percentage of Mammographic mass problem is 85%.

- Brest Cancer problem

This dataset is the study about patient who had undergone surgery for Brest Cancer. This task is to determine if the tumor is benign otherwise malignant.

Figure 3. Training results for Brest Cancer Problem

In Fig.3 it can be seen that in the training session of Brest cancer Wisconsin problem there is much cleared classification between two classes. There is no overlapping between the benign and malignant in the results.

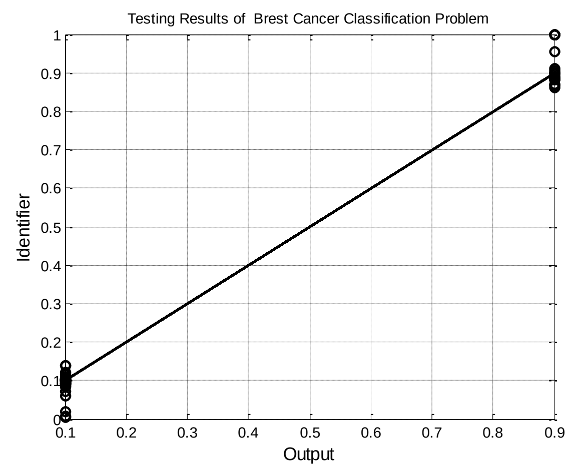

Figure 4. Testing results for Brest Cancer Problem.

In Fig 4. as outcome reveals from the study the testing results are same as training results in case of Brest cancer classification problem. The classification is much cleared.

- Mammographic mass Problem

Mammography is the most effective method for Brest cancer screening available. The Mammographic mass data come from the measurement of different parameters cancer suspected patient. This task is to determine of the bases of measured data patient benign or malignant.

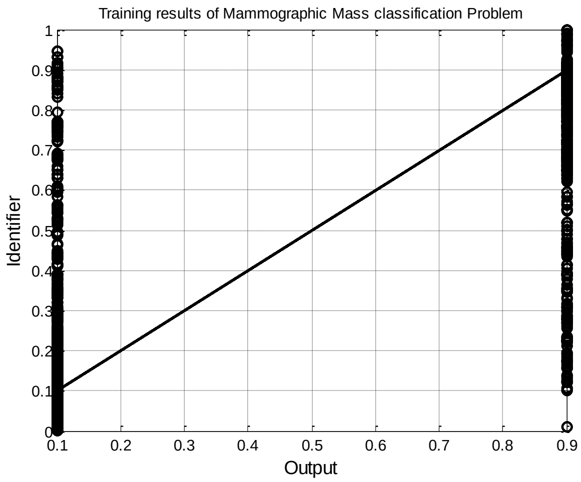

Figure 5. Training results for Cancer Problem.

As reveals from training results of mammographic problem in the Fig.5 that classification between benign and malignant is not very cleared. The both classes are overlapping each other. The MLP model could not learn the pattern of mammographic mass classification properly. The miss classification rate is maximum 15% in comparison to others problems.

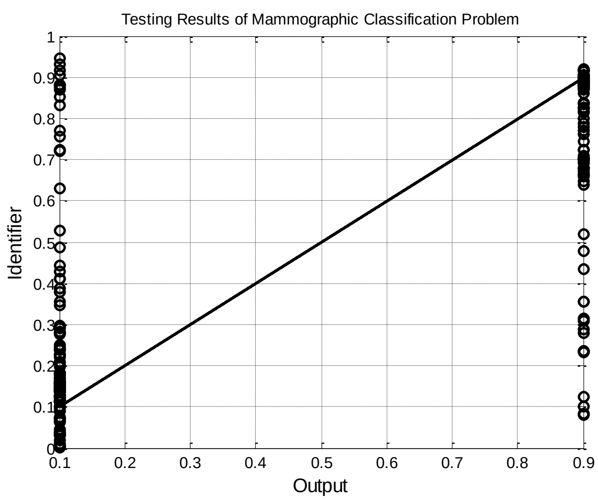

Figure 6. Testing results for Mammographic Mass Problem

In the testing we take the small subset of same data, the results are same as training of cancer classification problem as depicted in the Fig.6.

Table.1

Comparison of training and testing performance of two class problem

S.No.	Parameter	Brest Cancer	Cancer
1	Training goel, in term of MSE (error check)	0.00001	0.00001
2	Iteration needed	3000	3000
3	Training time in seconds	241	150
4	testing time in seconds	0.029	0.012
5	MSE for training data	0.0097	0.0334
6	MSE for testing data	0.0094	0.0406
7	No. of Attributes	10	6
8	No. of instants	699	961
9	Correlation coefficient training	0.9953	0.7659
10	Correlation coefficient testing	0.9956	0.7141
11	percentage of miss classification	0%	4.5%
12	learning late (Ƞ)	2.3	2.1
13	Numbers of hidden neurons	18	18

B. Three classes classification problems

The wine problem and the Iris classification problems having three types of classes. After observing the graph of MSE vs Iteration in Fig.2 it can be analyzed that red line of Iris problem minimized the error very quickly but still the MSE value of Iris is greater than wine MSE. On the other hand yellow line of wine takes time to minimize the error but minimize the error better way than Iris problem. It means that in the given data pattern the MLP model learn better for wine problem that iris problem. From the table 2 it can be seen that training correlation coefficient of wine is greater than that of Iris coefficient, The classification percentage of wine is 100% were classification percentage of Iris is 97%.

- Wine Problem

Wine dataset is using chemical analysis determine the origin of wine. There three origin of the wine as depicted in the Fig. 7 there are three classes of the wine.

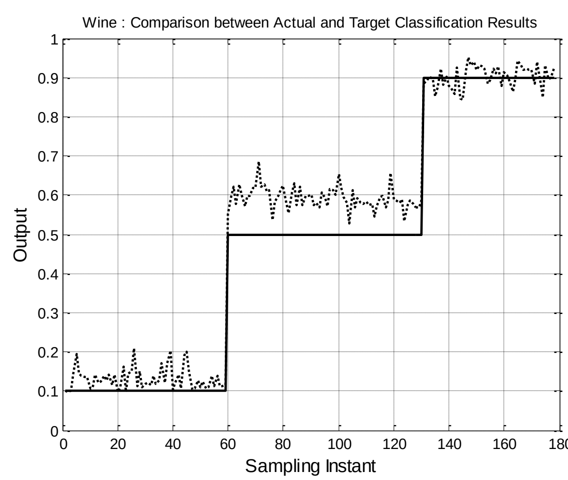

Figure. 7 Mean square error vs. iteration of training for IRIS problem

As depicted in the Fig.2 and Fig. 7 it can be observed, model has easily adopted wine classification problem. The classification performed by model better way.

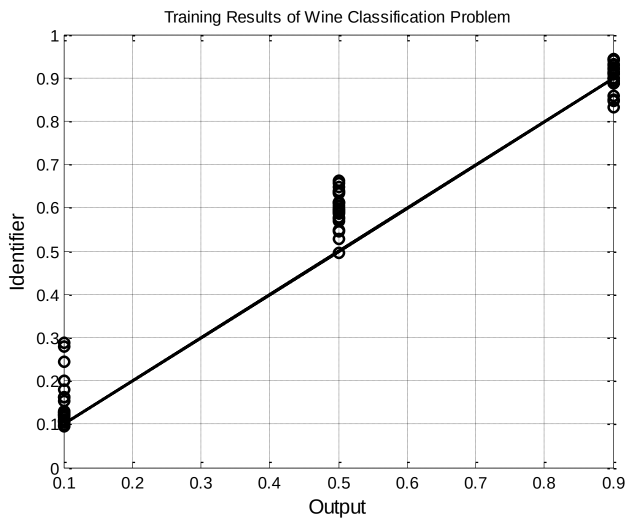

Figure 8. Training results for Brest Cancer Problem

The data of wine classification problem clearly classified. There is no overlapping between classes as depicted in the training results of wine problem.

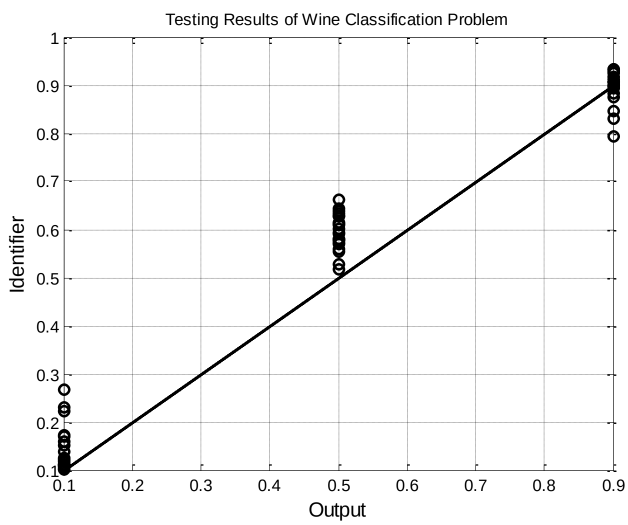

Figure 9. Testing results for Brest Cancer Problem

Same results are got by the testing observation (Fig. 9). In three class classification problem the wine classification results are better than Iris problem.

- Iris Problem

Iris data set is very popular dataset among researchers as Fisher Iris. It is open for all at university of California archive, 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.

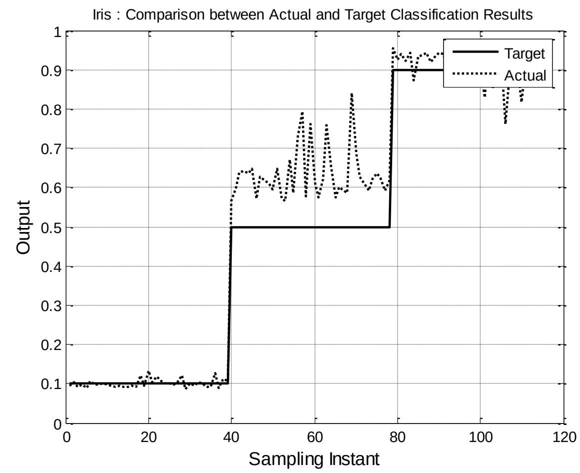

Figure 10. Mean square error vs. iteration of training for IRIS problem

From the Fig.2 and Fig. 10 it reveals that model adopts Iris classification problem in better way but in comparison of wine problem the performance of Model in context of iris poorer. The classified data of iris is more deviated than wine data.

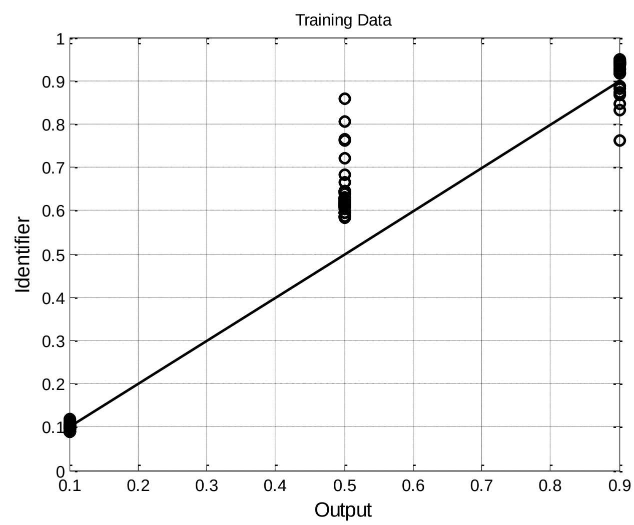

Figure 12. Training results for Iris Problem.

As the training results graph of Iris Fig.12 shows that upper two classes are overlapped with each others, that is undesirable for any classification problem. Even though the classification percentage of Iris is 97%.

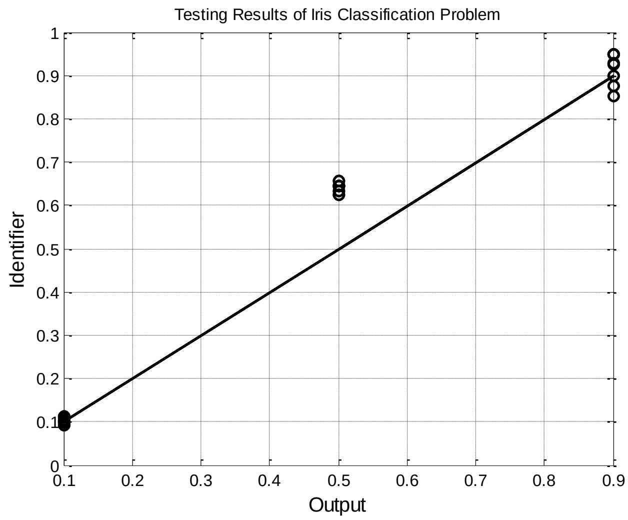

Figure 13. Testing results for Iris Problem

Testing results are good of model in context of Iris dataset as shown in the Fig. 13. There is cleared classified output data.

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 training and testing performance of three class problem

S.No.	Parameter	Wine	Iris
1	Training goel, in term of MSE (error check)	0.0001	0.00001
2	Iteration needed	3000	3000
3	Training time in seconds	115	105
4	testing time in seconds	0.0	0.0
5	MSE for training data	0.0022	0.0046
6	MSE for testing data	0.0025	0.0022
7	No. of Attributes	13	4
8	No. of instants	178	150
9	Correlation coefficient training	0.9890	0.9745
10	Correlation coefficient testing	0.9853	0.9869
11	percentage of miss classification	0%	3%
12	learning late (Ƞ)	0.66	2.1
13	Numbers of hidden neurons	24	36

V. Conclusion and Future Scope

It is to convey that in the entire classification problems, the best results have been taken with constant iteration (3000) for observation. Either the problem having two classes or three classes it is found in the observation that data patterns are significant without focusing on instants of data. It is found that Multiplicative Neuron model can learn properly in less instants of dataset. The quality of the data like noisy, missing value and range of data, standard deviation etc, and homogenous attributes affects the learning of Multilayer Neuron Model. It does not matter that how many attribute are contained by a dataset. The mammographic mass classification problem having 6 attributes where as Brest cancer classification have 10 attributes, against the perception the Brest cancer dataset having more attributes and complex mathematic learns better than mammographic mass problem.

This study provides the direction and scope for future work which include, role of hidden layer in Multilayer neuron model using different datasets. In the same way other model like Multiplicative, spiking and Integrate-and-fire Neuron model can be used for behavior analysis.

Acknowledgement

The authors thanks to Dr. William H. Wolberg, Wisconsin Hospitals Madison for Brest Cancer Dataset, Dr. Rudiger S. Wendtland for Mammographic Mass dataset, Michael Marshal for Iris dataset and Forina M. et al for wine dataset and special thanks to U.C.I. archives for getting available these datasets.

References

[1]Dan W Patterson "Inroduction 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] Peter Dayan and L F Abbott, Theoretical Neuroscience “Computational and Mathematical modeling of Neural System, MIT Press Cambridge, London.
[18] Pankaj K. Kandpal, Ashish Mehta, comparison analysis of single Multiplicative neuron with conventional neuron model, IJET,660-666(2017), 2249-3255.
[19] Pankaj K. Kandpal, Ashish Mehta, comparison and analysis of Multiplicative neuron and Multilayer Perceptrons using three different datasets,IJERGS, Vol. 05, issue 03, 2017
[20] Pankaj K. Kandpal, Ashish Mehta, compatitive analysis of different Arificial neuron models,IJSAE, VOL. 05, Issue 07, Jul-2017.

Authors Profile

Mr. Pankaj Kumar Kandpal pursuied Bachelor in Computer Application from IGNOU University New Delhi, India in 2006 and Master of Computer Application from IGNOU University New Delhi, India in 2007 (integrated). He is currently pursuing Ph.D. from Kumaon University, Nainital, Uttarakhand, India and currently working as JAO in BSNL, an Inian Government PSU. He was a member of “Computer Society of India” reputed journal in India. His main research work focuses on Artificial Intelligence and Machine Learning. He has eight years of telecommunicartion industries experience.

