Multilayer Perceptron Networks

KoliPoojaVasant
Department of Electronics
And telecommunication
Rajarambapu institute of
Technology, Islampur
poojakoli7194@gmail.com

AwatiJayshree S.
Department of Electronics
And telecommunication
Rajarambapu institute of
Technology, Islampur
jayashree.samai@ritindia.edu

Abstract

The paper focuses on the various neural network techniques consisting of Multilayer perceptron (MLP) neural community. The diverse components of the neural networks techniques are noted inside the paper at the side of the advantages and some of the drawbacks. Neural networks are taken the different strategies to the trouble fixing than that of the conventional computer systems. The neural network fashions are trained with measured values of the sphere power at the arbitrary factors of the network. Additionally the paper researched some essential issues regarding with the functionality of the multilayer perceptron's with one or hidden layer. This result may be very useful inside the analyzing pattern of diverse varieties of popularity and database retrieval. The neural networks are also beneficial inside the photo processing for image recognition.

Keywords*: Multilayer Perceptrons,Neural networks.*

INTRODUCTION

Neural networks are taken theunique approaches to the problem solving than that of the traditional computer systems. Traditional computers use an algorithmic procedures i.e. computer follows the set of instructions within the given order to resolve the problems. till the precise steps that are the laptop wishes to observe the known laptop can not resolve a problem. However computer systems could be the a lot beneficial if they do the matters that don't exactly know a way to do the matters. The Neural networks

Process information in the same way that like a human brain does. The neural network composed of the huge quantity of highly interconnected neurons running inside the parallel to clear up the precise problem. They may be now not programmed to perform the specific mission. On the alternative side, the traditional computer systems are use a cognitive approach to solve the trouble and the way to clear up the problem is to solved must be regarded and said within the small unambiguous commands. These commands are then transformed to the excessive stage language software and then into the system code for information of the computer that could understand the technique. those machines are completely predictable manner all of us can expect if anything is going incorrect is because of the hassle of software program.

The neural networks and the conventional algorithmic computers are the complement of every other no longer in the competition. The neural networks are the normally divided within the terms in their related schooling algorithms that is unsupervised network, constant-weights network, and the supervised network. There may be no getting to know required for the fixed-weight networks consequently the mastering mode is supervised or an unmanaged. The supervised gaining knowledge of networks has the primary move of the neural community model development. The schooling statistics encompass many pairs of the inputs and outputs of training styles. So for the unsupervised learning rule the training set includes the handiest the input education patterns. So the community is educated with out the benefit of any instructor.

The network learned to the adaptand based on the experiences that collected by the previous training patterns.

NEURAL NETWORK

An Artificial Neural Network (ANN) is an information of processing the paradigm which are stimulated by way of the manner of organic anxious systems, such that the mind and their method records. A fundamental element of the paradigm is the radical shape of the statistics processing of the machine. And it's miles composed of the massive quantity of exceptionally interconnected processing elements that operating within the unison to clear up the specific issues. The ANN is designed for the precise programs such that the sample popularity,data classification by through the learning processes.

A. Multilayer Perceptron (MLP) network

The common part of the neural network model is multilayer perceptron (MLP). The aim of this type of the network is to create the model that is correctly measure the input to the output using the historical data and so that an Analytical Study of Neural network. The model can be used to produce the output when the desired output is known to programmer. A graphical representation of the multilayer perceptronis showed as below:

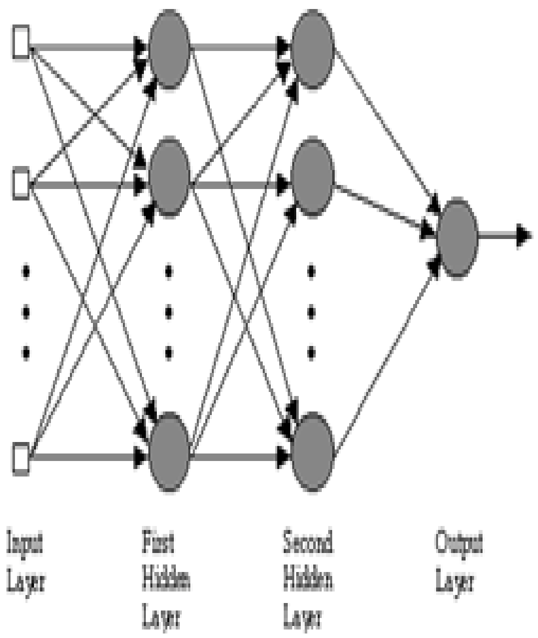

Fig 1.*Architecture of MLP neural network*

The class of the neural networks includes the multiple layers and the computational units, the interconnected neurons in the feed-forward type of way. Every neuron in one of the layer has been directed connections to neurons of a related layer. The multi-layer networks are uses variety of the learning techniques such that the most popular is the back-propagation algorithm. Sothat the output values are compared to the correct answer to compute the value of some of the predefined error-functions. By the various types of techniques, the error is then fed back through the network. And using the information that the algorithm adjusts the weights of each connection in order to decrease the value of error function by some of small amount. In case the one can say that the network has beenlearned the certain target function which determined. So that to adjust the weights of the network properly anyone applies the general common method for the nonlinear optimization that are called the gradient descent of the network. The derivative function of the error function with respect to the network weights has calculated and therefore the weights haschanged such that the error will be decrease. For the reason the back-propagation algorithm can only be applied to the networks with the differentiable activation functions

Table 1*-Case Processing Summary*

		N	Percent
Sample	Training	42	73.7%
	Testing	15	26.3%
Valid		57	100.0%
Excluded		0	
Total		57	

Table 1 describes the processing data as follows training data is 73.7% and testing percent is 26.3% and case processing data is 100%.

Table 2:*Network Information*

Input Layer	Factors	1	h1
		2	h2
	Covariates	1	y1
		2	y2
	Number of Units[a]		6
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers		2
	Number of Units in Hidden Layer 1[a]		3
	Number of Units in Hidden Layer 2[a]		2
	Activation Function		Sigmoid
Output Layer	Dependent Variables	1	x1
		2	x2
		3	x3
	Number of Units		4
	Rescaling Method for Scale Dependents		Normalized
	Activation Function		Sigmoid
	Error Function		Sum of Squares

Table 2:*Network Information*

Input Layer	Factors	1	h1
		2	h2
	Covariates	1	y1
		2	y2
	Number of Units[a]		6
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers		2
	Number of Units in Hidden Layer 1[a]		3
	Number of Units in Hidden Layer 2[a]		2
	Activation Function		Sigmoid
Output Layer	Dependent Variables	1	x1
		2	x2
		3	x3
	Number of Units		4
	Rescaling Method for Scale Dependents		Normalized
	Activation Function		Sigmoid
	Error Function		Sum of Squares

a. Excluding the bias unit

Above table describes network information. Network consists of input layer, hidden layer and output layer. Y1 and Y2 are the dependent variables and X1, X2 and x3 are covariates of the network shown in fig 2. Tan sigmoid transfer function is used.

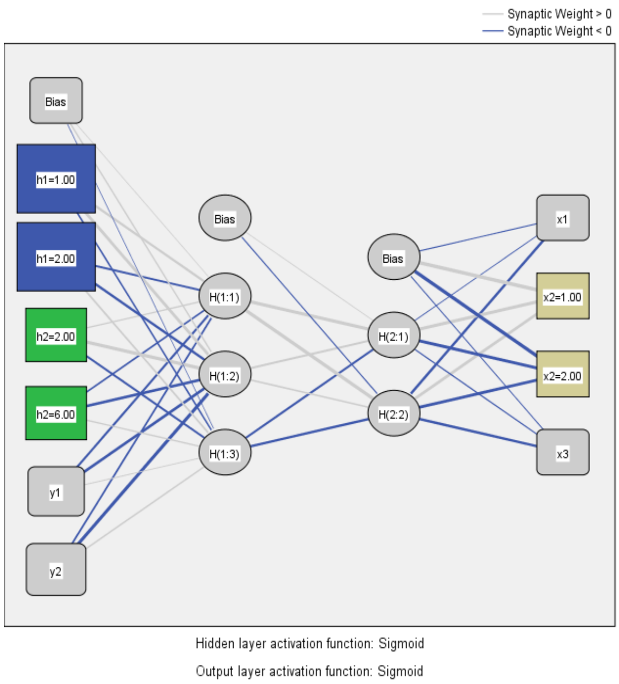

Fig 2: network Diagram

Above fig shows the network diagram with x1 and x2 nodes in input layer, h1 and h2 is hidden layer node and y1 and y2 are the output nodes.

Table 3: Model summary

Model Summary

Training	Sum of Squares Error	2.027
	Average Overall Relative Error	.869
	Percent Incorrect Predictions for Categorical Dependents x2	2.4%
	Relative Error for Scale Dependents x1	.776
	Relative Error for Scale Dependents x3	.768
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	00:00:00.010
Testing	Sum of Squares Error	.211
	Average Overall Relative Error	.956
	Percent Incorrect Predictions for Categorical Dependents x2	.0%
	Relative Error for Scale Dependents x1	.953
	Relative Error for Scale Dependents x3	.959

a. Error computations are based on the testing sample.

Model summary for network shown in fig 2 is divided into two sections training and testing. Training data gives sun of square error value as 21.85, and testing data gives 3.038 as sum of square error

Table 4: Parameter estimates

Parameter Estimates

Predictor		Predicted								
		Hidden Layer 1			Hidden Layer 2		Output Layer			
		H(1:1)	H(1:2)	H(1:3)	H(2:1)	H(2:2)	x1	[x2=1.00]	[x2=2.00]	x3
Input Layer	(Bias)	.039	.607	-.017						
	[h1=1.00]	1.504	2.066	-.802						
	[h1=2.00]	-1.055	-1.669	.999						
	[h2=2.00]	.551	2.448	-1.193						
	[h2=6.00]	-.864	-1.788	.564						
	y1	-1.393	-1.834	.305						
	y2	-1.006	-3.132	.683						
Hidden Layer 1	(Bias)				.216	-.423				
	H(1:1)				2.323	3.475				
	H(1:2)				1.187	.699				
	H(1:3)				-1.220	-1.649				
Hidden Layer 2	(Bias)						-.325	4.038	-3.480	-.306
	H(2:1)						-.296	2.299	-2.069	-.464
	H(2:2)						-1.700	1.838	-1.993	-1.677

Table 5: classification

Classification

Sample	Observed	Predicted		
		1	2	Percent Correct
Training	1	41	0	100.0%
	2	1	0	.0%
	Overall Percent	100.0%	.0%	97.6%
Testing	1	15	0	100.0%
	2	0	0	.0%
	Overall Percent	100.0%	.0%	100.0%

Dependent Variable: x2

Above table shows the training and testing classification data. Training data is 100% and testing data is 100%.

Predicted by Observed Charts

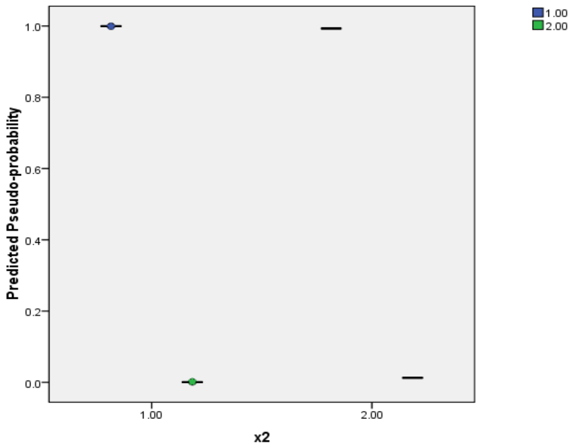

Fig 3: *Predicted Y1 chart by observer*

The ROC curve is graph of sensitivity and specificity.

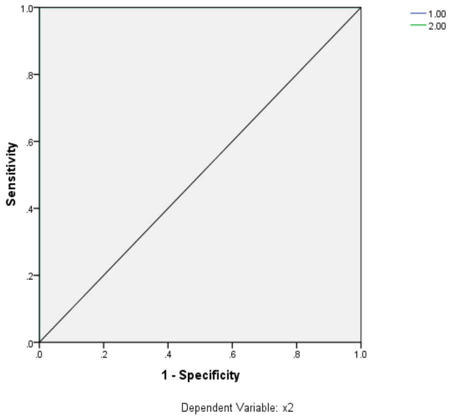

Fig 4: *ROC curve*

Table 5: *Area under the curve*

Area Under the Curve

		Area
x2	1	1.000
	2	1.000

Above table contains area under curve of x2 node and its value is 1.

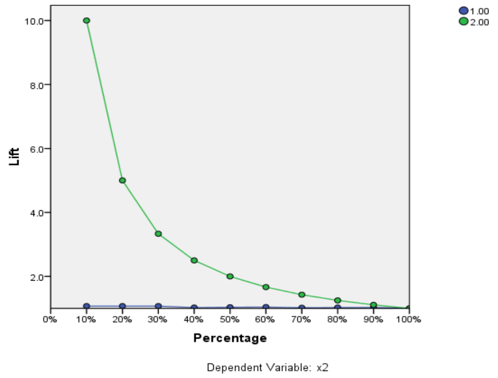

Fig 5: *Lift chart*

Above fig shows lift chart for dependent variable y1between 10 – 100 % lift value is 10.

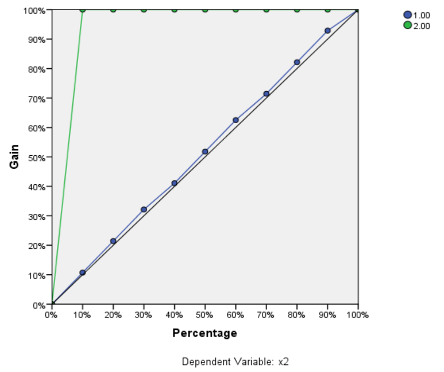

Fig 6:*Cummulative gain chart*

Cumulative gain chart is graph of gain versus percentage. It is observed that gain 10% up to 100% and then increases with percentage

Table6: *Independent variable importance*

	Importance	Normalized Importance
h1	.347	100.0%
h2	.230	66.4%
y1	.200	57.8%
y2	.222	64.1%

Above table shows importance of variables h1 has 100% importance, h2 has 66.4% and y1 has 57.8%, y2 has 64.1% importance.

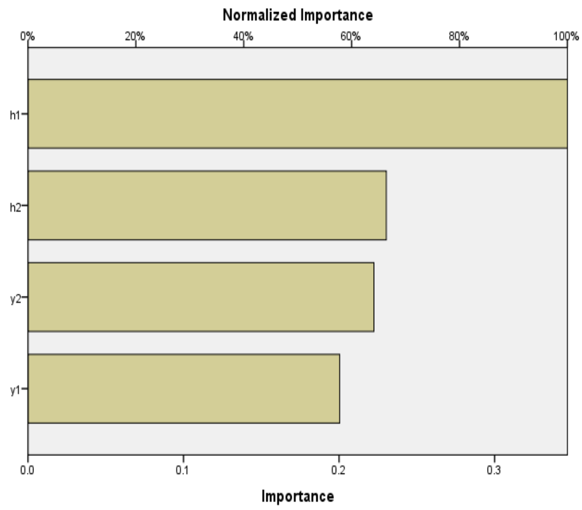

Fig 7: *Independent variable importance*

Above graph shows importance of variables h1 has 100% importance and h2 has 65% importance, y2 has 60% importance and y1 has 58% importance

CONCLUSIONS

The design of the multilayer perceptron network designed to analyze the facts that the MLP with more number of hidden layers gives the better performance than the one or two number of hidden layers specifically. The design regarding to the output performance of the parameters of the layers of the network.Also, the difference between the output of the multilayer perceptron and the mathematical formula that defines the output performance of the parameter that are best suite with the experimental procedure. And for the multilayer neurons there are various tables shown above that give detail information related to the multilayer perceptions and neural network. So for the future work the recommend to use the hybrid approaches, which are the combination of an artificial neural network with the other techniques like expert systems of the Fuzzy logic and the Genetic Algorithm (GA) to make such type of analysis.

REFERENCES

1. Bitzer S. and Kiebel J. 2012. Recognizing Recurrent Neural Networks (rRNN): Bayesian Inference for Recurrent Neural Networks. arXiv: 1201, 4339v1.
2. Calcagno G. and StaianoAntonino. 2010. A Multilayer Neural Network-based approach for the identification of responsiveness to interferon therapy in multiple sclerosis patients. Information Sciences. Vol. 180, Issue 21, 41534163.
3. Hill T, O'Connor M, Remus W. Neural Networks Models for Time Series Forecasts, Management Sciences, doi: 10.1287/mnsc.42.7.1082 1996; 42:1082-1092.
4. Hann TH, Steurer E. Much ado about nothing? Exchange rate forecasting: Neural networks vs. linear models using monthly and weekly data. Neurocomputing 1996; 10:323339.
5. K. Fukushima. Neocognitron: A self-organizing neural network model for a

mechanism of pattern recognition unaffected by shift in position. Biological cybernetics, 36(4):193–202, 1980.

6. V. Jain and E. Learned-Miller. Fddb: A benchmark for face detection in unconstrained settings. Technical Report UM-CS-2010-009, University of Massachusetts, Amherst, 2010.