Time Series Analysis of Forecasting Indian Rainfall

Akashdeep Gupta, Anjali Gautam, Chirag Jain, Himanshu Prasad, Neeta Verma

***Abstract*—This paper presents a study of neural network model for prediction of Indian rainfall. The purpose of this paper is to evaluate the applicability of ANN. In this paper the performance of different networks have been evaluated and tested.The multilayered artificial neural network with learning by back-propagation algorithm is used .The paper implements weather prediction by building training and testing data sets and finding the number of hidden neurons in these layers for the best performance. The proposed model has been able to predict values with suitable results. The prediction is made on the bases of previous data. The criteria for prediction in the model are correlation, RMSE, standard deviation .Prediction of Rainfall is necessary for Agricultural & Metrological Department. In India, most of our Economy is dependent on agriculture. A big percentage of GDP is contributed by agriculture. In India, agriculture provides around 70% of employment either directly or indirectly. This is major reason for analysis of prediction of rainfall.**

***Keywords*- Artificial Neural Network, Root Mean Square Error, Standard Deviation, and Backpropagation.**

I. INTRODUCTION

Prediction of monsoon rainfall is the necessity of the metrology department. In a country like India where around 70% of the total population is dependent on the agriculture directly or indirectly prediction of rainfall plays a very major role.

In this study we are predicting the summer monsoon rainfall using the Artificial Neural Network (ANN) based on the time series analysis. ANNs can be used for classification, pattern recognition, function approximation and forecasting. Before ANN and other AI models, these tasks were carried out by statistical methods such as the linear and nonlinear regression, principal components analysis and the Bayesian classification models, but it is observed that the ANNs model is better than the linear regression models.

In the present study we use a multi layer feed forward Artificial Neural Network. A multilayer feed forward network typically consists of interconnected set of processing units called the neurons. These neurons are categorized into three distinct layers: the Input layer, Output layer and the Hidden layer. Each link between the Input layer and Hidden layer or the link between Hidden layer and the Output layer has a weight associated with it.

Manuscript Received on May 2013.

Akashdeep Gupta, Inderprastha Engineering College, Ghaziabad (U.P.), India
Anjali Gautam, Inderprastha Engineering College, Ghaziabad (U.P.), India
Chirag Jain, Inderprastha Engineering College, Ghaziabad (U.P.), India
Himanshu Prasad, Inderprastha Engineering College, Ghaziabad (U.P.), India
Neeta Verma, Inderprastha Engineering College, Ghaziabad (U.P.), India

Hidden layers are necessary to capture the non linear dependencies between our data features and the variable we are trying to predict. Choosing the number of hidden layers in a network depends on the training and validation data. Each neuron in the network performs two operations:

1. It makes a weighted sum of its input from the input layer.

2. Then, it transfers the weighted sum to its output layer using its activation function.

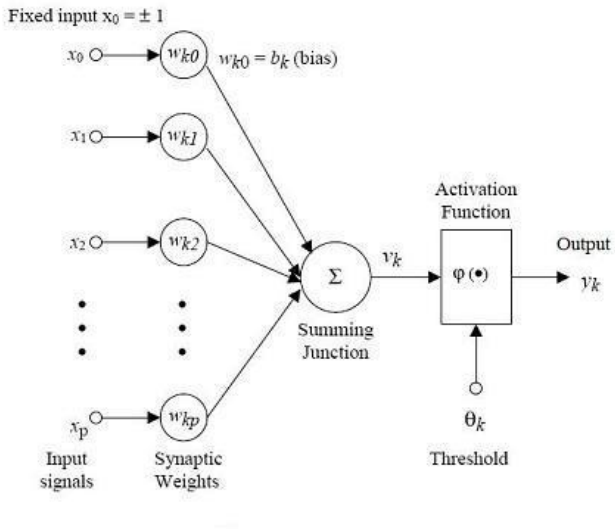

Fig.1. Describes the Architecture of ANN.

II. DATA COLLECTION

Selection of Dataset is important in order to achieve the objective in this study.

The 140 year monthly data set of (1871-2010) of ALL-INDIA RAINFALL of 30 meteorological subdivisions encompassing 2,880,324 SQ. KM. with a resolution of up to 0.1 mm/month obtained from the Indian Institute of Tropical Meteorology website with the original source as referred by the department being the Indian Meteorological Department (IMD), shall be used for the analysis.

III. DATA PREPROCESSING

1. The input and the output data obtained have to be normalized because they are of different units and otherwise there will beno correlation between the input and the output values.We normalize the collected data using Formulae in range [0.2, 0.8] .

$$Xn = [(X - Xmin)/(Xmax - Xmin)] * 0.6 + 0.2$$

2. After obtaining the normalized data, the next step is to train the input data using Matlab Back-propagation Algorithm. The proposed ANN model is basically a three layered ANN back propagation learning. Total samples are 1506 .Out of total samples (: 92% (130 yr data) are used for training Rest 8% (10 yr data) are used for kept for testing.

(a) Training: 92% (130 yr data)
(b) Testing: 8% (10 yr data)

IV. METHODOLGY

In order to perform the weather forecasting using neural network and comparing the performance of neural network model with different learning rates and by setting different number of neurons in hidden layer, we are using backpropagation algorithm and supervised learning.
The following Steps are followed in the process:-
1. Data preprocessing.
2. Defining the ANN model.
3. Training of ANN.
4. Testing of data

5. Comparing the results.

Network Model: In our study the ANN model consists one hidden layer and an output layer. One neuron for input and one for the output. In this model we are using four neurons for the hidden layer with a learning rate of 0.2.

Backpropagation: Backpropagation is a common method of training artificial neural networks so as to minimize the objective function. It is a supervised learning method, and is a generalization of the delta rule. It requires a dataset of the desired output for many inputs, making up the training set. It is most useful for feed-forward networks (networks that have no feedback, or simply, that have no connections that loop). The term is an abbreviation for "backward propagation of errors". Backpropagation requires that the activation function used by the artificial neurons be differentiable.
In this study we use two Activation function for
1. Hidden neuron is hyperbolic tangent (tanh(x))
2. Output neuron is linear

V. EXPERIMENTAL RESULT.

In this experiment, the weather dataset is implemented into ANN in order to investigate the output. The performance is measured in terms of its RMSE (Root Mean Square Error), correlation. The correlation and RMSE is calculated and compared. The comparison shows the successful ANN model is developed.
Initially an ANN model with two neurons in hidden layer & a learning rate 0.2 is defined. The RMSE of the target & output is obtained and compared with standard deviation (the RMSE obtained should be less than the Standard deviation).The analysis of all Neural network with different Neuron (upto 5) in hidden layer is done.
The best network that has least RMSE (< SD) & more correlation b/w target & output is considered for testing purpose.
The best network is used for testing new data set (with the left 10 years data). The RMSE of target & output is calculated and compared with standard deviation.

Training analysis

S. no	No. of neuron	epoch	Correlation Between Input & Target	Correlation Between Target & Output	RMSE	Standard Deviation (SD)
1	2	1000	0.7487	0.7404	0.1114	0.165
2	3	1000	0.7487	0.7557	0.1086	0.165
3	4	1000	0.7487	0.7608	0.1069	0.165
4	5	1000	0.7487	0.7562	0.1084	0.165

Fig 2 describes the comparative study of training data using ANN model with different neurons.

Testing Result

S.no	No. of neuron	Correlation B/W Target & Output	RMSE	SD
1	4	0.7476	0.0529	.1708
2	5	0.7315	0.1239	.1708

Figure 3

Fig 3 describes the testing results.

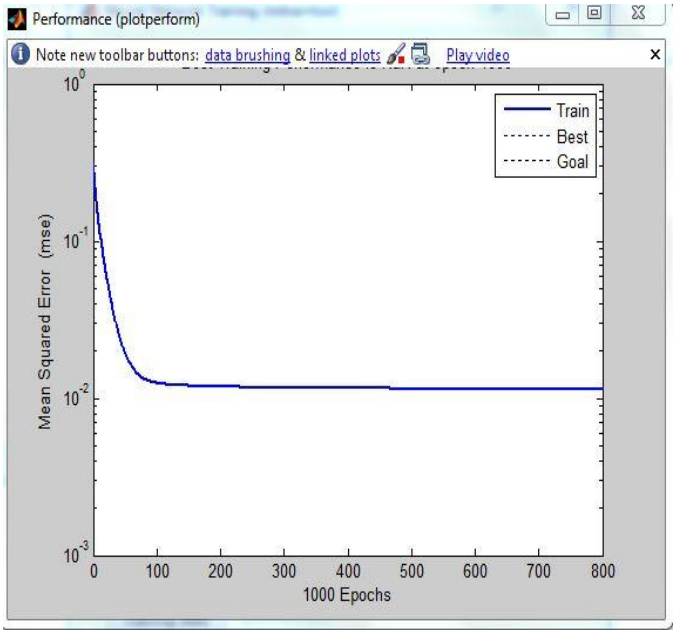

Fig. 4 Performance graph with 4 neuron in hidden layer

The best ANN model with 4 neuron in hidden layer is considered for testing with this the less RMSE & correlation is more. Using this model prediction of Rainfall is done using Time series analysis

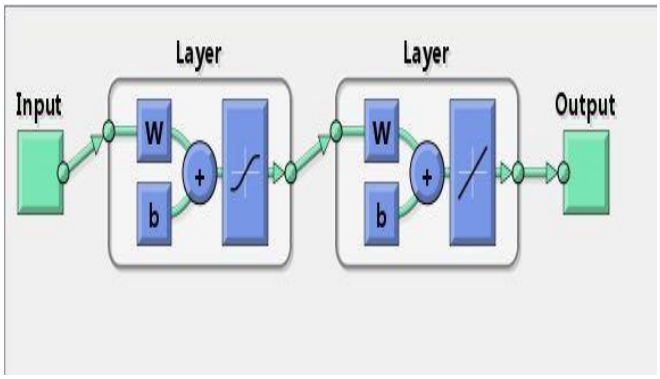

Fig.5 Architecture of neural network

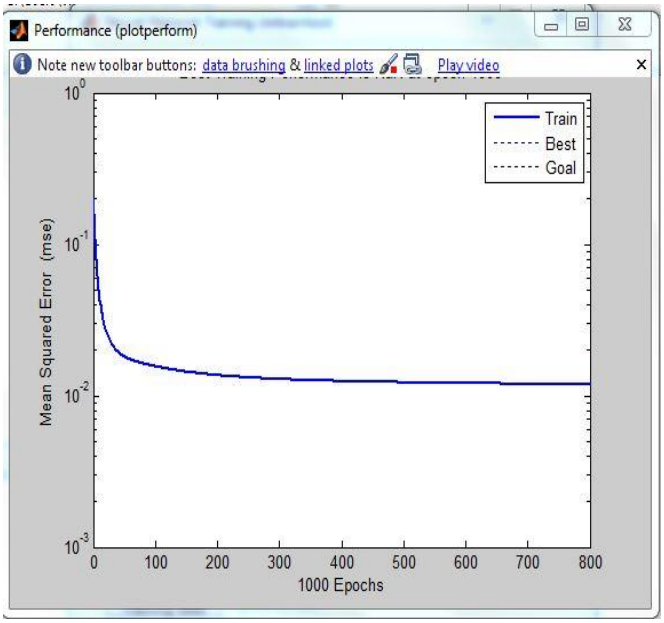

Fig. 6 Performance graph with 5 neuron in hidden layer

VI. CONCLUSION

For rainfall prediction, Artificial Neural Network was applied. According to the results backpropagation neural network were acceptably accurate and can be used for predicting the rainfall. The designed ANN model can be used for prediction of Indian rainfall. Using the prescribed model the prediction of values is closer to the actual values. The conclusion with this study is that with the increase of hidden neurons first the error decreases up to a certain limit and then it starts increasing because the noise gets fitted in the network. More number of hidden neurons in the network represents more non-linearity in the data. In our study we started with two neurons and goes up to 5 neurons and the best model was of 4 hidden neurons which gives the least Root Mean Square Error (RMSE).

VII ACKNOWLEDGEMENT

The authors would like to thanks to Dr. K.C Tripathi for valuable time, data, and information of neural network & describe various techniques of neural network tool of Matlab.

References

[1] Tripathi ,K.C. ,Das ,I.M.L , and Sahai ,A.K. “Predictability of sea surface temperature anamolies in the Indian ocean using Artifical neural network” ,Indian Journal of Marine Sciences ,vol.35(3), September 2006 ,pp. 210-220.

[2] Tripathi ,K.C. ,Das ,I.M.L , and Panday ,A.C. , ”Southern Indian Ocean indices as early predictors of Indian summer monsoon ” ,Indian Journal of Marine Science vol. 37(1) ,March 2008 ,pp. 70-76

[3] Hsieh W. William ,Tang Benyang “Applying neural network models to prediction and data analysis in meterology and oceanography “Bulletin of the American Meteorological Society.

[4] Paras, Sanjay Mathur, Avinash Kumar, and Mahesh Chandra, "A feature based on weather prediction using ANN"World Academy of Science, Engineering and Technology 34 2007

[5] (ftp://www.tropmet.res.in/pub/data/rain/iitm-regionrf.txt).

[6] Enireddy Vamsidhar et. al.,"Prediction of rainfall Using Backpropagation Neural Network Model", International Journal on Computer Science and Engineering Vol. 02, No. 04, 2010, 1119-1121

[7]]Dr S. Santosh Baboo and I. Khadar Shareef, "An efficient Weather Forecasting Model using Artificial Neural Network", International Journal of Environmental Science and Development, Vol. 1, No. 4,October 2010.

[8] Ben Krose and Patrick van der Smagt , "An introduction to neuralnetworks", Eighth edition, November 1996.