Air Temperature Prediction using Artificial Neural Network for Anyigba, North-Central Nigeria

Akolo J. A.
Centre for Atmospheric Research
Kogi State University Campus
Anyigba Kogi State Nigeria

Felix A. S.
Department of Computer Science
Federal Polytechnic, Idah
Kogi State Nigeria

Okoh D.
Centre for Atmospheric Research
Kogi State University Campus
Anyigba, Kogi State Nigeria

ABSTRACT
Temperature prediction is the application of science and technology to predict the state of the temperature for a future time and a given location. Temperature prediction is important to protect life and property. Temperature prediction is made by collecting quantitative data about the current state of the atmosphere. In this paper, an artificial neural network was employed using the Levenberg-Marquardt backpropagation to develop the prediction model. The Root Mean Square Error is then calculated between the perceptron and the desired output for the input vectors. This back-propagation approach is chosen for this training because, from the works of literature reviewed, it is regarded as one of the most efficient training algorithms because of its fast and stable convergence and suitable for training small and medium-sized problems in the artificial neural-networks field. Our study used four years' data (2013-2016) gotten from Atmospheric Monitoring Equipment Network Automatic weather station situated at Centre for Atmospheric Research, Anyigba North Central Nigeria. The proposed model is tested using the network of the hidden neuron with the least root mean square error to the observed data. The outcome of the predicted values is compared with the observed values. After comparing the results with the observed values, it shows that our model has the potential for temperature prediction.

General Terms
Artificial Neural Network, Air Temperature, Back propagation.

Keywords
Artificial Neural Network, Backpropagation, Multi-layer perceptron, Air Temperature Prediction.

1. INTRODUCTION
Temperature prediction is the application of science and technology to predict the state of the air temperature for a future time and a given location. An attempt has been made to predict air temperature since ancient times [1]. In recent times, air temperature prediction is made by collecting data about the current state of the atmosphere using scientific understanding to project how the temperature will evolve in the future. Air temperature prediction might not stop extremely high or cold air temperatures, breakdown in operational equipment, low agricultural yields or even spread of diseases associated with high or low air temperature, but helps us in taking timely necessary actions and precautions. Air Temperature prediction can be carried out using the following steps.

- Data collection (Temperature)
- Data assimilation and analysis
- Numerical weather prediction
- Model output postprocessing

A neural network is a powerful data modeling tool that can capture and represent complex input /output relationships. The Artificial Neural Network (ANN) is a machine learning prediction algorithm that is inspired by the animal nervous system. In recent literature, ANN is used to predict the network traffic pattern. ANN is considered as it is not a complex algorithm for hardware implementation [3]. Artificial Neural Networks (ANNs) have been used in various domains for modeling and prediction with high accuracy due to its ability to learn and adapt [5][6]. Neural networks can generalize and are resistant to noise. Neural networks have seen an explosion of interest over the years, and are being successfully applied across an extraordinary range of problems [8]. Indeed, anywhere that there are problems of prediction, classification or control, neural networks are being introduced. The Levenberg - Marquardt Backpropagation Learning Algorithm is a kind of supervised learning algorithm used in a multilayer perceptron. In this work, it requires three layers, which are the input, hidden and output layers to train the multilayer perceptron. There are two phases of the back-propagation algorithm; the feed-forward and error back-propagation. The feed-forward propagates the input vectors (variables), that is; Year, Day of the year and hour of the day through the layers to produce the output vector (variable), which in

2. DATA COLLECTION
The air temperature data of four years (2013-2016) was considered in this work and obtained from the data archive of Atmospheric Monitoring Equipment Network (AMEN) at the Centre for Atmospheric Research situated at the Kogi State University Campus, Anyigba with latitude 7.4934^{o}N, and longitude 7.1736^{o}E with and a landmass of 29,833km North Central Nigeria. The data were measured using an HMP50 probe sensor, which is installed on the meteorological station, it records measurement and stored the data on the format of Year, Month Day Hour and Minutes (i.e. Five (5) minutes mean value) on a data logger, and the sample of the observed data collected is as shown in figure 1. The five minutes data was later averaged to hourly for this work. The sample of the normalized data after averaging is shown in figure 2.

u

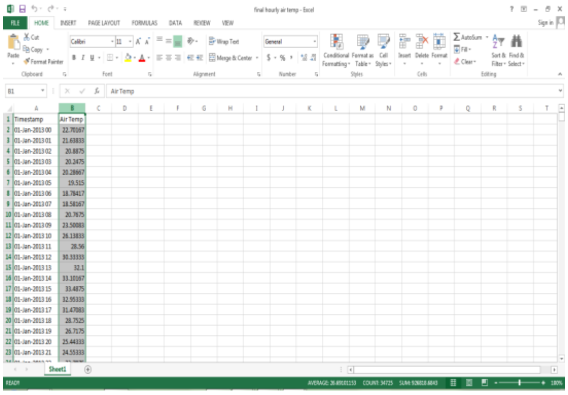
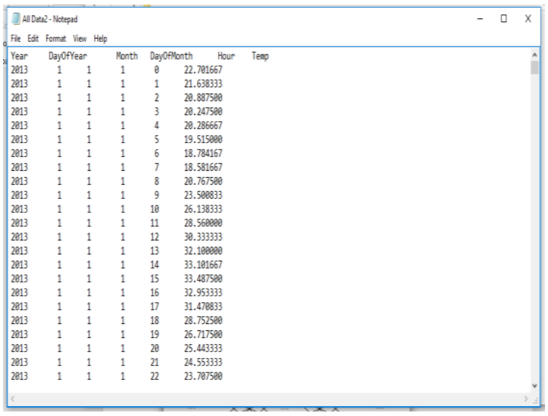


Fig. 1 Sample data from AMEN station Anyigba



Year	DayOfYear	Month	DayOfMonth	Hour	Temp
2013	1	1	1	0	22.701667
2013	1	1	1	1	21.638333
2013	1	1	1	2	20.887500
2013	1	1	1	3	20.247500
2013	1	1	1	4	20.286667
2013	1	1	1	5	19.515000
2013	1	1	1	6	18.784167
2013	1	1	1	7	18.581667
2013	1	1	1	8	20.767500
2013	1	1	1	9	23.500833
2013	1	1	1	10	26.138333
2013	1	1	1	11	28.560000
2013	1	1	1	12	30.333333
2013	1	1	1	13	32.100000
2013	1	1	1	14	33.101667
2013	1	1	1	15	33.487500
2013	1	1	1	16	32.953333
2013	1	1	1	17	31.470833
2013	1	1	1	18	28.752500
2013	1	1	1	19	26.717500
2013	1	1	1	20	25.443333
2013	1	1	1	21	24.553333
2013	1	1	1	22	23.707500

Fig.2 Sample of Normalized data

The chosen atmospheric data were divided into two randomly selected groups, the training group, corresponding to 70% of the patterns, and the test group, corresponding to 30% of patterns; so that the generalization capacity of the network could be checked after the training phase.

The Root-Mean-Square Error (RMSE) is used to measure the difference between values (sample and population values) predicted by a model or an estimator and the values observed. In other words, the RMSE represents the sample standard deviation of the differences between predicted values and observed values. The RMSE was used as the basis of determining the degree of precision of this model.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (P_i - O_i)^2}{n}}$$

Where Oi is the observed values, Pi is the predicted values and n is the number of all observed values used in each case.

The general structure of the input/outputs i.e. the ANN model for temperature prediction is as shown in Figure 3.

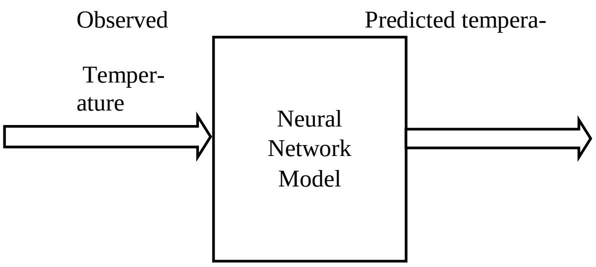


Fig. 3 General Structure

3. ARTIFICIAL NEURAL NETWORKS (ANN)

Artificial Neural networks could be defined as an interconnected of simple processing elements whose functionality is based on the biological neuron. The biological neuron is a unique piece of equipment that carries information or a bit of knowledge and transfers to other neurons in the chain of networks. An artificial neural network is also a computational technique that is designed to simulate how the human brain processes a specific task by a huge distributed parallel processing [9]. A neural network is made up of simple processing units. These units are called neurons. Neurons are mathematical elements which have a nervous property in that they store practical knowledge and empirical information to make it available to the user and that by adjusting the weights [3][8][9]. The ANN model has been designed by using four basic numbers of procedures: (1) Data collection, (2) Data processing, (3) building the Artificial Neural Network, (4) training, validating and testing network as shown in figure 4 [9].

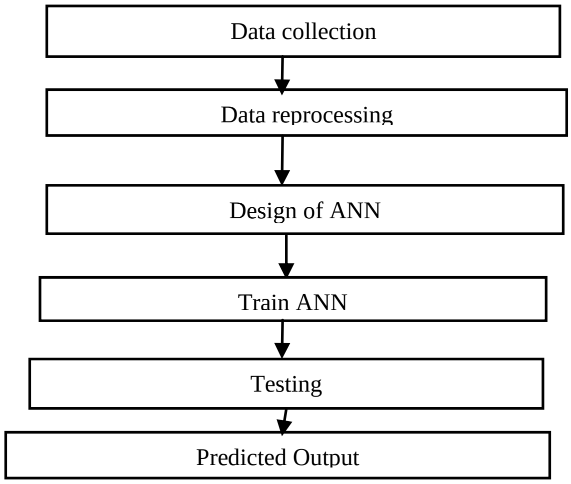


Fig. 4 The block diagram of the training procedure

The developed ANN model is based on one of the neural network architectures named multi-layer perceptron.

3.1 Multi-Layer Perceptron

A multilayer perceptron (MLP) can be defined as a class of feedforward artificial neural network. It is a feedforward artificial neural network that generates outputs from a set of inputs. MLP has at least three layers of nodes: an input layer, a hidden layer, and an output layer. Each node is a neuron that uses a nonlinear activation function apart from the input node. An MLP uses backpropagation as a supervised learning technique. Since there are multiple layers of neurons, MLP is a deep learning technique [12]. It is applicable in the field of computational neuroscience and parallel distributed processing, speech recognition, image recognition, and machine translation.

In multilayer perceptron design, the most important issues are specifying the number of hidden layers and the number units in these layers [4][9][10]. Once the number of layers and the Pnumber of units in each layer has been selected, the network's weights and thresholds must be set to minimize the prediction error made by the network. An example of a multilayer perceptron is shown in figure 4 [13]

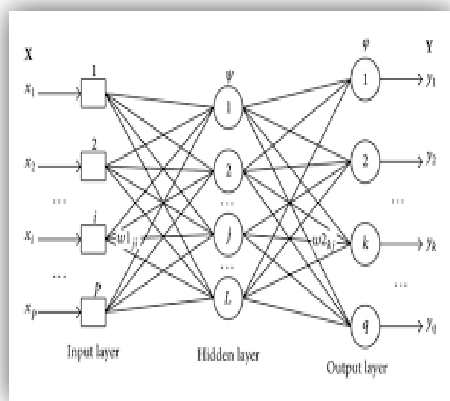


Fig. 4 Example of Multilayer Perceptron

3.2 The Levenberg-Marquart Algorithm

The Levenberg-Marquart backpropagation algorithm is the best-known example of a neural network training algorithm. The Levenberg-Marquardt is one of the fastest algorithms. Levenberg-Marquardt Algorithm (LMA) is the most widely used optimization algorithm. It provides numerical solutions to the problem of minimizing a nonlinear function [2][4][9].

The LMA which is a blend of the steepest descent method and the Gauss-Newton (GN) Algorithm inherits the stability property of the steepest descent method and the speed advantage of the Gauss-Newton algorithm. This accounts for why the LMA can converge well even if the error surface is much more complex than the quadratic situation. Hence, it is more robust than the GNA. The basic idea of the Levenberg-Marquardt algorithm is that it performs a combined training process around the area with complex curvature, the Levenberg-Marquardt algorithm switches to the steepest descent algorithm until the local curvature is proper to make a quadratic approximation, then it approximately becomes the Gauss-Newton algorithm which can speed up the convergence significantly.

The Levenberg-Marquardt algorithm can be derived in four different ways i.e. steepest descent algorithm, Newton's method, Gauss-Newton algorithm, and the Levenberg-Marquardt Algorithm. In this, work we look at the Levenberg-Marquardt Algorithm

3.3 Working Principle of the LMA in a Three-Layer ANNs Architecture.

The Levenberg - Marquardt Backpropagation Learning Algorithm is a kind of supervised learning algorithm used in a multilayer perceptron. In this work, it requires three layers, which are the input, hidden and output layers to train the multilayer perceptron. There are two phases of the back-propagation algorithm; the feed-forward and error back-propagation. The feed-forward propagates the input vectors (variables), that is; Year, Day of the year and hour of the day through the layers to produce the output vector (variable), which in this case, is the hourly air temperature, in degree Celsius (^oC), for Anyigba. The Root Mean Square Error is then calculated between the perceptron and the desired output for the input vectors. The error back-propagation involves propagating the error back through the layers from output to hidden and input layers. In each layer and each neuron in the layer, the synaptic weights are updated. This process is iterated for all the input vectors until the perceptron is converged to the solution. This can be determined from the root mean square error of each input vector. We are using the error-correction learning rule whereby the neural network is trained with input data for which some desired output is known. The difference between the actual neural network output and the desired output is called an error (or error information). This error is then used to find the network error function. The objective of the error-correction learning rule is to minimize the error function (i.e. minimize the difference between the network output and the desired output, which is an optimization problem). This back-propagation approach is chosen for this training because from the kinds of literature reviewed, it is regarded as one of the most efficient training algorithms because of its fast and stable convergence and suitable for training small and medium-sized

4. INSTRUMENTATION

The instrument used to capture data for this work is the Campbell Scientific CR1000 datalogger. The standard station is a fully configured, solar-powered, automated weather station. It consists of a weatherproof enclosure that contains a highly reliable Campbell Scientific datalogger, 12V battery and charge controller. The weather station is equipped with a standard set of sensors that takes records of meteorological parameters including the HMP50 probe for measuring air temperature measured in degree Celsius (C). The data logger is programmed using CR basic for the sensors supplied when completely connected the weather station will automatically start to take measurements through each of the parameter sensors outside of the box. Additional sensors especially dual-sensor can be added as options. The HMP50 probe measures temperature with a 1000 Ohm platinum resistance thermometer (PRT) and relative humidity (RH) with a 50Y Inter-cap capacitive chip. The chip is field-replaceable, which eliminates the downtime typically required for the recalibration process. This temperature and RH sensor are compatible with all Campbell Scientific dataloggers. The HMP50 should be housed in a solar radiation shield typically the 41303-5A. The 41303-5A 6-plate naturally aspirated shield attaches to a mast, cross arm, or tower leg. To attach the radiation, shield directly to a tripod mast, tower mast, or tower leg, place the u-bolt in the side holes. To attach the 41303-5A to a CM202, CM204, or CM206 cross arm, place the 41303-5A's u-bolt in the bottom holes. The HMP50 probe needs a supply voltage of 7 to 28 Vdc (typically powered by the data logger's 12v supply). It has a consumption rate of 2mA and a length of 2.8inch (7.1cm) and a diameter of 0.47inch (1.2cm). It has a measurement range of -40^{o}C to +60^{o}C. Figure 5 is a typical view of the installed station with an enclosed box containing the datalogger.

Fig. 5. Campbell Scientific Automatic Weather Station

While figure 6 shows a typical HMP50 probe sensor installed in the Automatic weather station.

Fig.6. HMP50 Probe

5. RESULTS AND DISCUSSION

The optimal structures for developed Artificial Neural Network for obtaining minimum prediction error are shown in Table I. The network with the least RMSE should be used as the predicting model.

Table 1. ANN model structure

Number of hidden layers	1
Number of hidden neurons	50
Number of epochs	100
Activation function used in inputs layer	Linear
Activation function used in hidden layers	Sigmoid
Activation function used in output layers	Sigmoid
Learning method	Supervised

The Average Root-Mean-Square error RMSE is shown in figure 7

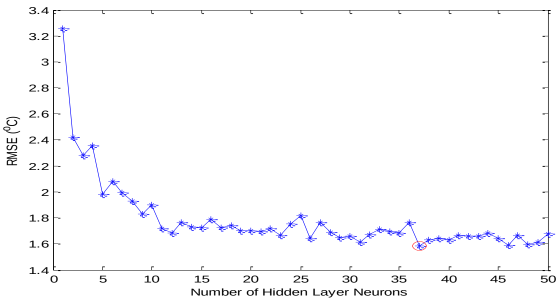


Fig. 7 Average RMSE

As it can be seen from the Figure above, the best optimized' numbers of hidden layers for the network is 37 number of neurons, to stress it more we can generally say that among the 50 architectures, the architecture with 37 neurons in the hidden layer was chosen in this study because it has the lowest RMSE of about 1.49^{O}C and therefore considered the optimum point. A larger or smaller number of neurons in a hidden layer will result in over-fitting and under-fitting respectively.

Table 2. Shows the total RMSE of the selected days compared with the average RMSE of the selected days and the average RMSE of the ANN

Selected days	RMSE	Average RMSE	Total RMSE
07-Jan-2013	0.80	34	1.49
20-Jan-2014	0.70	35	1.49
30-Dec-205	0.70	34	1.49
20-Feb-2016	0.73	36.5	1.49

Table 3. Shows observed and predicted values for the selected days

Selected days	Observed value	Predicted value
07-Jan-2013	34	34
20-Jan-2014	35,5	35
30-Dec-205	34	34
20-Feb-2016	36	36.5

The table above shows the observed and the predicted values from the selected days so that the generalization capacity of the network can be checked after the training and the testing phase.

Table 4. shows the minimum and maximum temperature of the selected days for the observed values

Selected days	Minimum temperature	Maximum temperature
07-Jan-2013	16	34
20-Jan-2014	23	35.5
30-Dec-205	18.5	34
20-Feb-2016	26.5	36

Table V. shows the minimum and maximum temperature of the selected days for the predicted values

Selected days	Minimum temperature	Maximum temperature
07-Jan-2013	18	34
20-Jan-2014	22	35
30-Dec-205	19	34
20-Feb-2016	26	36.5

5.1 Graphical Comparison of the Observed and Predicted Values

The observed and predicted values for each selected day by Artificial Neural Network are shown in Fig.8 (a-d). From the result shown in Fig.7, and Tables 2, 3, 4, and 5. The model developed has been applied to randomly selected days and years. i.e. January 2013, January 2014, December 2015 and February 2016 in that order. The outputs from the model for the selected days of the years were plotted alongside the Observations values recorded from the Campbell scientific weather station of AMEN in Anyigba. This is aimed at establishing the degree of accuracy of the predictions made by the model and hence validate the possibility of using this model in atmospheric/meteorological parameter predictions especially for the fact that there are few or no available atmospheric/meteorological stations in many locations in Nigeria. Generally, it was observed that the RMSE value of the selected days differs from each other for the fact that the air temperature from the observation varies on daily basis and the model predicted values follow the trends of each day. This simply means the model is in agreement with the observation and can be used to predict such variations. The RMSE obtained for the days considered shows good agreement between the model predictions and the observations. As expected, the transitions in the Neural Network model predictions are smoother than in the physical observations since models include smoothing operations in learning the transitions between observations. Also, the figures expectedly show that the maximum air temperature of about 36 to 38 degree Celsius in each day are recorded at about 3 to 4 pm local time, corresponding to the peak period in this region, except in few cases where there are sudden changes in the measurement values of the air temperature. This sudden change is attributed to the cloud covers which serve as a shield of the incoming solar irradiance by either scattering or absorption. The air temperature starts decreasing at around 18^{th} to 19th hour of each day (corresponding to the period of sunset in Anyigba), where it eventually gets to little above 22^{O} C and around the 0 to 1 hour when it gets to as low as 18^{O}C as the Sun’s disappeared completely from our horizon.

It is revealed from the figures that the model predictions are very good and close to the measurements. The ranges of the predictions are generally smaller than those of the observations. This is expected since model predictions are more tilted to averages in the data pattern to give smoother transitions.

In comparison between the selected days for the years considered in this work, the highest RMSE is 0.80^{O}C and the least is 0.70^{O}C. Overall, statistics of the errors (model output-observation differences) show that the model runs at an average range of 0.7325^{O}C in error for future prediction. Also, Using the Root Mean Square Error (RMSE) as a means of comparison, the result shows that the model predictions are in good agreement with the values having an average Root Mean Square Error of 0.7325^{O}C. This result of validation and the comparative study indicates that the Neural Networks based technique is suitable for air temperature prediction. This study, therefore confirms the ability of artificial neural networks to predicts air temperatures for Anyigba. It is observed that the predicted values are in good agreement with observed values and the predicted error is very less. Therefore, the proposed ANN model with the developed structure shown in Table I can perform good prediction with the least error.

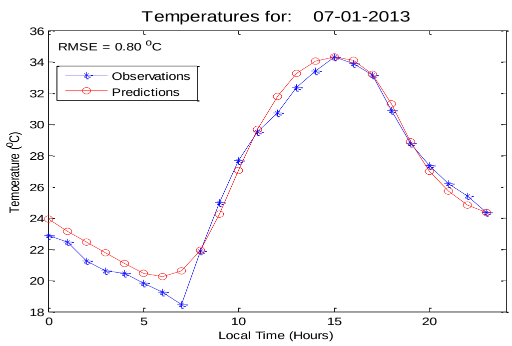


Fig. 8(a) Observations/Predictions for 07th January, 2013.

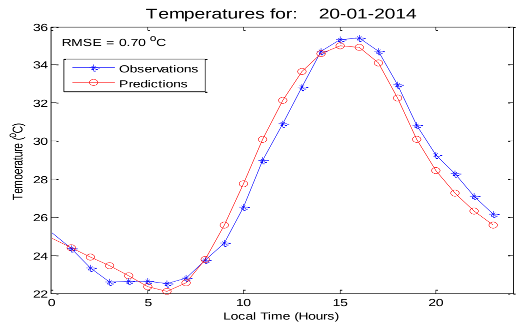


Fig. 9(b) Observations/Predictions for 20th January, 2014.

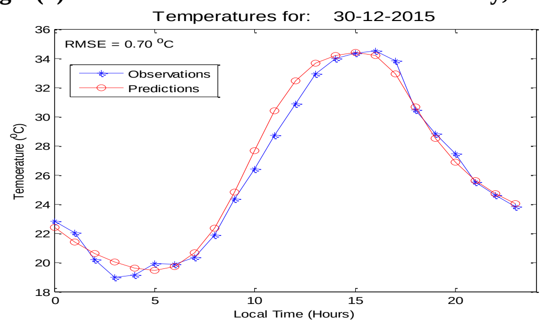


Fig 8(c) Observations/Predictions for 30th December, 2015

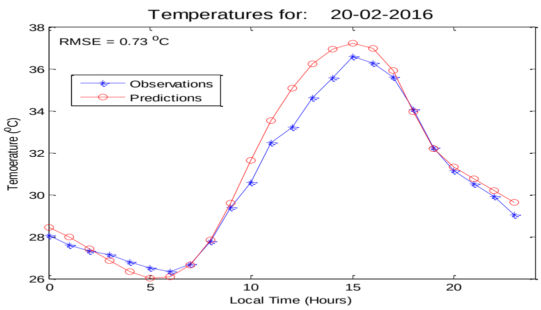


Fig. 8(d) Observations/Predictions for 20th February, 2016.

6. CONCLUSION

It is established at the end of this work that the ANN model’s prediction is in good agreement with the observation of the stations and therefore can simulate an in-situ observation practically in our location; this means ANN is applicable in meteorological parameters prediction in a tropical environment. The model is therefore suitable for short term prediction of air temperature in Anyigba for which air temperature is required for agriculture, health-related issues, academia as well as for smooth operations of heavy industrial machines.

7. ACKNOWLEDGMENTS

We would like to express our gratitude to Prof. A.B Rabiu, the Director/CEO and Mr. Najib Yusuf the Head Research and development division of the Centre for Atmospheric Research, Kogi State University Campus, Anyigba North-Central Nigeria for providing us the required data. Special thanks to all staff of the Research and Development Division.

8. REFERENCES

[1] Brian A. Smith (2001). *Air Temperature Prediction using Artificial Neural Network*. A master thesis submitted to the University of Georgia, Athens

[2] Ananth Ranganathan (2004). *The Levenberg-Marquardt Algorithm*. 8th June, 2004

[3] C. Gershenson "*Artificial Neural Networks for Beginners*" Networks,2003., vol. cs.NE/0308,

[4] J. Shao (1997). *Improving Nowcast of Road surface Temperature by a backpropagation Neural Networks*. A weather and forecasting journal Vol 13, October 1997

[5] Haykin, S. (2008). *Neural Networks: A Comprehensive Foundation*, Prentice Hall, ISBN 0-13-273350-1

[6] Haykin S. (1994). *Neural Networks: A comprehensive foundation NY*: Macmillan p.2

[7] D. B Shank, Gerrit Hoongerboom, R.W McClendon (2008*). Dewpoint temperature prediction using Artificial Neural Networks*. Journal for Applied meteorology and climatology. June 2008

[8] Mohsen Hayati and Zahra Mohebi (2007). *Application of ANN for temperature forecasting*. World academy of science and technology. International journal of electrical and computing engineering. Vol 1 No. 4. 2007

[9] Manal A. Ashour, Somia A. Elzahaby and Mahmoud I. Abdalla (2016). *Backpropagation neural Neural Network approach for mean temperature prediction.* www.arpapress.com/Volumes/Vol29Issue1/IJRRAS_29_1_02.pdf 12. **IJRRAS** 29 (1). October 2016

[10] Kumar Abhishek. M.P Singh, Suswata Ghosh and Abhishek Annand (2012). *Weather forecasting model using Artificial neural networks.* December 2012

[11] M.T. Hagan. H.B. Demuth, M.H. Beale. *Neural Network Design*. PWS Publishing Company, Boston, Massachusetts, 1996.

[12] D. Patterson. *Artificial neural networks*. Singapore, Prentice Hall. Good wide-ranging coverage of topics, although less detailed than some other books, 1996.

[13] www. Skymind.ai

[14] www.wikipedia.com