Artificial Neural Network based Weather Prediction using Back Propagation Technique

Saboor Ahmad Kakar[1],
Naveed Sheikh[2]
Department of Mathematics
University of Balochistan Quetta

Saleem Iqbal[4], Abdul Rehman[5],
Aziz ullah Kakar[6]
Department of Mathematics
University of Balochistan Quetta

Hazrat Ali Kakar[8]
Department of Mathematics
University of Balochistan
Quetta

Adnan Naseem[3]
Department of Computer Science
COMSATS University Islamabad

Bilal Ahmad Kakar[7]
Department of Computer Science
University of Balochistan Quetta

Bilal Khan[9]
Department of Computer Science
COMSATS University Islamabad

Abstract—Weather forecasting is a natural phenomenon which has some chaotic changes happening with the passage of time. It has become an essential topic of research due to some abrupt scenarios of weather. As the data of forecast is nonlinear and follows some irregular trends and patterns, there are many traditional techniques (the literature like nonlinear statistics) to work on the efficiency of models to make prediction better than previous models. However, Artificial Neural Network (ANN) has so far evolved out to be as a better way to improve the accuracy and reliability. The ANN is one of the most fastest growing technique of machine learning considered as non-linear predictive models to perform classification and prediction weather forecasts maximum temperature for the whole days (365) of the year. Therefore, a multi-layered neural network is designed and trained with the existing dataset and obtained a relationship between the existing non-linear parameters of weather. Eleven weather features were used to perform classification of weather into four types. Furthermore, twenty training examples from 1997-2015 were used to predict eleven weather features. The results revealed that by increasing the number of hidden layers, the trained neural network can classify and predict the weather variables with less error.

Keywords—Weather forecasting; artificial neural network; classification; prediction; backpropagation; hidden layers

I. INTRODUCTION

Due to sudden changes happens in the nature and weather conditions, the massive amount of power is always required to manage and keep the balance for the atmosphere. Weather forecasts the natural phenomena, involved the vague understanding of atmosphere processes due to the time difference. Weather forecast is a process of rapidly changing conditions like humidity, rainfall, wind speed, directions, temperature, etc. It is used to collect this data with some defined tools like, meteorological satellites, radars, wireless sensor and high-speed computers to perform computation. The data which is recorded through satellites are always available in the form of an images to get the whole idea of abrupt changes happened in the environment. Apparently, weather is an intensive, chaotic and dynamic process makes it a fierce challenge.

As whole it is considered, as an important matter to be focused on all over the world, because the weather prediction can have expressive impact on variant parts of the world. This prediction activity is particularly handled by individuals to improve the efficiency of operations performed on daily basis. It is one of the most essential and in demand responsibility taking over by meteorological services. And all the decisions are taken under the uncertain situation associated with local and global climate variables. Different researchers and scientists are performing research since 1920's to predict the disasters and changes in climate conditions. The most important factor that varies is time series. It involves a combination of observation, trends and patterns both for regional and global. So far, researchers are trying to predict the weather forecasting by different means which increase the computational power. As we are talking about prediction which means to determine something going to happen in future. In a past few decades, the excessive increment in a computational power has forced to design and develop the complex systems having the ability to deal with the hardness of power systems to perform calculations. Machine learning and data mining [13, 14, 16] are commonly used fields for this research as these are actively used to improve the computational power. For weather forecast most of the systems depends on conditions to adjust to prevent the loss or damage to some extent.

The weather forecast systems need to be intelligent that they can easily read the statistical data to generate patterns and rules to study and based on past data predict the future. Recent study [15] has shown that numerical weather prediction model could be introduced to represent the global importance of forecast. They simulated the sub grid lakes in global forecasts for reduction at high latitude in forecast errors especially in spring and summer. Although, above described tools like radars, wireless sensors and satellites are more enough for handling conditions of weather but these improvements and enhancement are result of betterment for future time. And the better observation under computational techniques. Since, many researchers have made efforts to identify the well-defined forecasting models which includes linear, non-linear and hybrid techniques to improve accuracy. It has been

observed from recent studies that in some cases the hybrid computational models [17,18] results in decomposition of linear and nonlinear forms and considered as efficient approach comparative to single or individual techniques. It also has less number of chances to produce errors. On contrary some of studies claims that hybrid or combination of models not proved to be efficient. Such kind of uncertainties in this case always open the door for new opportunities to make models more precise. Even due to high penalty of imbalance systems relying always on meteorological forecast is not a good option.

This research is focused to propose an efficient model based on ANN to improve the computational power to make future prediction more efficient. The selected features to perform the experiments for actual weather is sunny, foggy, thunderstorm and rainy for the course period of 20 years (7000 days).

The study is categorized as: section II represents the study done previously, section III show the methodology has been used for this research, section IV demonstrated the results along with discussion and at last conclusion of the study is presented.

II. LITERATURE REVIEW

Most of the work have been done on different weather forecasting applications by using prediction techniques especially in machine learning field. Some of the work is summarized in this section.

Authors in [1] focused on reduction of random and numerical errors with the post processing of Numerical weather predictions (NWP) models to propose the best error free model which will be used for weather forecasts purposes. They had used most inspiring the Kalman filter to perform analog forecast in an ordered form. Along with it, adapted the weighing strategy to verify the forecast analog for specific time and location. With this new approach, they have successfully achieved the improvement in comparison of other algorithms. With the strong combination of both, NWP altogether and analog by producing the hybrid of these two models with logistic regression as a preprocessing [2]. Authors had cover up the probabilistically 2 m weather forecast with their efficient hybrid technique as shown to outperform comparatively others. Somehow this is not performing good for 10 m speed (wind). They considered, the representative NWP for long term purpose to overcome the wind speed factors involved like ambient temperature, atmospheric pressure, local terrain, etc. [3] Authors included the machine learning by joining the neural networks with NWP to assess the wind speed which offered a satisfying improvement in performance. They have proposed the new method with the use of feed forward artificial neural network for weather prediction which is proved to be a better approach as compared to traditional approaches [4]. Different datamining techniques has also been used for classification of weather parameters such as, the C5 decision tree with the consideration of some basic factors like min and max temperature, wind speed and rainfall prediction for a specific month or year [5]. These techniques been observed in the months by giving enough data to influence the weather prediction with some expected changes in identified climate patterns. Authors used a combination of TLFN neural network

and RNN architectures to make ANN models through observe pattern variables to detect the relationships for future predictions. At the same time, their prediction analysis was performed with linear regression to achieve the better results of prediction with great number of accuracy [8]. Multi classification is proved to be great in their research by considering the large number of features. According to given results, their model is particularly providing an efficient prediction. The classification techniques are a possible solution of severe occurrences like socio economic or environmental [6]. However, the large number of meteorological data makes it difficult or somehow impossible to perform analysis for weather prediction. By considering the classification approach of ANN we can conclude that their approach is far better than traditional to identify associated patterns for consecutive events.

Consecutively, there are more complex strategies to leverage the load forecasting by using hierarchy patterns in distributions network for small area [7]. It is used for parameter optimization to make innovative forecast. Authors proposed a new approach named Dilation erosion perceptron with a fusion of modified genetic algorithm which is based on a complete theory of mathematical morphology for weather prediction [9]. They have also considered to remove the temporal distortions in weather forecasting by proposing a method the automatic phase fix procedure. And their evolutionary process model demonstrated the good results. In [10] somehow, we find it difficult to figure out the prediction patterns because of some consistencies happened in characteristics of weather. Authors targeted the occurrence of changes specifically conducting in a huge region of Bangladesh. Their model shows reliable prediction for this seasonal change. Classification is commonly encountered as a more powerful decision-making idea in ANN [11]. Even appeared as an evolutionary task in data mining and machine learning fields. Neural Networks are being applied for different mathematical and optimization purposes such as forecasting, image processing, object recognition, classification, prediction, processing of signals, etc. Authors considered an important factor to perform some operations for solar online power forecasting, with the selection of advanced statistical methods. Their online solar power application got trained for operational planning of system. One of the benefit could be noticed of ANN based photovoltaic system is that it could handle the abrupt circumstances or variations by this system. Proposed radial based function network makes prediction for 24 hours. They have proved that given model worked with 90% reliability for crystal clear sky on the other hand 80% for unclear sky condition as the weather uncertainties always led to great errors. In this research, the authors proposed two ANN based approached one of them is used d to predict the rainfall rate for 2001 to 2013 period through sensitivity analysis [12]. However only one model is proved to be efficient as compared to other one.

III. METHODOLOGY

This section introduces the methodology being used and the analysis performed. In this research, basically two things are considered; weather classification and weather prediction. Further, the weather data is classified into four types thunderstorm, Rainy, Foggy, and sunny then predicts weather

features for next year after training on 20 years on ANN. And the 20 training examples are the data acquired for the year 1997 – 2015. There are 11 prediction features have been taken into account which are MaxDewpoint, MaxHumidity, MaxPressure, MaxTemperatue, MaxVisibility, MeanWindSpeed, MinDewpoint, MinHumidity, MinPressure, MinTemperatue, MinVisibility.

A. Artificial Neural Networks

ANN has been used since 1943 and inspired from human brain reactions. Just like human brain needs several neurons to transmit information and perform the actions just like that ANN needs neurons to perform the action in hidden layers. The general architectural diagram of ANN has been shown in Fig. 1. They are capable enough to train or learn more complex functions. It consists of input layer, hidden layer and output layer. The number of neurons is equal to the number of input variables present. The complexity of model always depends on the scenario to perform action iteratively. If you want your model to perform better and efficient activities, you need to pass it to large number of data and number of hidden layers. Then some activation functions like hyperbolic, tangent, sigmoid, SoftMax, etc. makes the work easier to accommodate the model they evaluate the values passed to it to compute error value. It has multiple types of neural networks. A neural network which is made up of processing units has a natural ability to store knowledge. Knowledge in network is acquired from the environment through given learning process. And the weights are the strengths of neurons to store the knowledge.
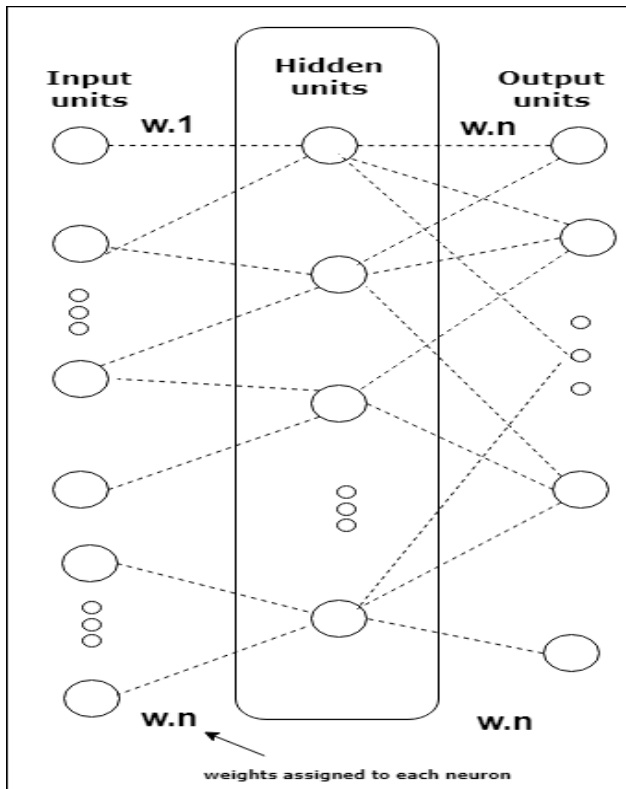
Fig. 1. General Architecture of ANN.

B. Architecture of ANN based Model

Our proposed architectural model is shown in Fig. 2. This diagram represents that increased number of neurons in hidden layers always raise the ability of network to perform efficiently by getting data from input nodes. It is also called as fully connected network; if every single node of one layer is connected to every other node in adjacent layer. The theta values in the network represents the weights as they stores the processing information of neurons attached with the layers. After training the data the network hidden layers conclude the desired output for a set of previously unseen input data. The model is closely prepared to process the information through neurons. The data is gathered from what the project predicted and what the actual weather was. The mathematically general model is explained step by step below. The parameters have been used in the algorithm are:

- ➢ backPropClassifier – has the classification algorithm through back propagation
- ➢ backPropPredictor – has the prediction algorithm through back propagation
- ➢ forwardPropClassifier – has the classification algorithm through forward propagation
- ➢ forwardPropPredictor – has the prediction algorithm through forward propagation
- ➢ Classifier – has the main classification algorithm.

1) Load a file and save it in a variable 'X'. All files of weather data are gathered from the period '1997-2015' are loaded.
2) Number of features (total 11 features+5 bias units) = Number of features (Max humidity, Mean wind speed,etc.) +bias units.
3) Y=X (:, Number of features]) →classes of weather are extracted from X; such as ': represents 16 column in Y.
4) X (:, Number of features]= [ ] → 16 column of X is emptied.
5) Weather classes in Y are bit mapped → they are bit mapped is because it is easy to match numerical output.
6) This is how we get [X Y] array. such as X is an input and Y is output.
7) Now we will define some variables: -

- • IEPSILON = 1 (weight) → Initial theta of each layer
- • LAMBDA = 0.01→ Regulation parameter
- • NUMBER OF FEATURES → Number of weather features
- • M → Number of training examples
- • K → Number of output classification
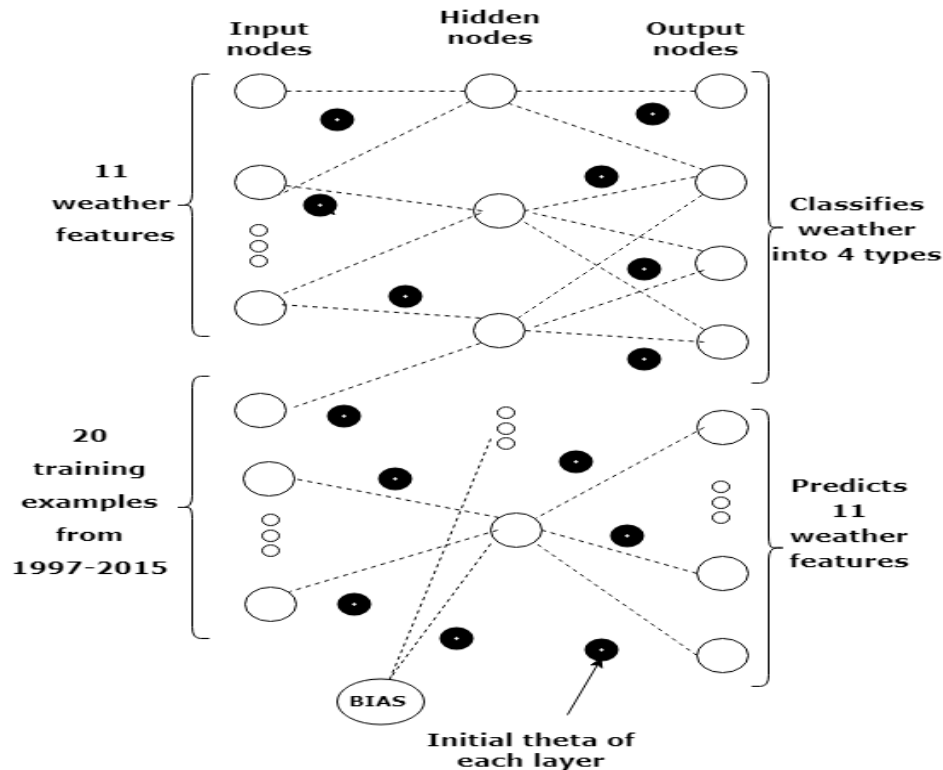- • m = 365 → DAYS IN A YEAR

Fig. 2. Proposed Model of ANN Architectural Diagram.

8) Then we choose initial theta matrix for layer 1 and layer 2 of Neural Network through formula: -

- ITheta1 = 2 * IEPSILON * rand (Number of features-1, Number of features) - IEPSILON * ones (Number of features - 1, Number of features)

9) 'Predictor' is called to train Neural Network:

- It takes array X, ITheta1, ITheta2, LAMBDA as input.

- Few variables are already defined assigned to new ones: -

    - E = M-m;

    - x = X(:, 1)

    - y = X(:, 366)

- A loop is executed to separate 365 values of a single feature.

- ITheta1 & ITheta2 are placed inside thetavec.

- Lost function is minimized through 'prediction cost function'

    - Optimset → create structures for optimization functions

    - Fminuc → tries to minimize a function.

- Steps:

1) Reshape thetavc & get ITheta1 and ITheta2 from it.
2) Apply forward propagation:
    1. Input x, theta1, theta2
    2. output a1, a2,a3
3) Computing error
    1. sqrError= sum(sum((a3-y). ^2))
    2. sqrError = $\sum(\sum(a3-y)^2)$
4) Computing cost function
    1. Jval = (1/(2*e)) * (sqrError+ LAMBDA*sum(thetavc.*2))
    2. Jval = 1/2e(sqrError+LAMBDA($\sum$ thetavc^2))
5) Apply backpropagation through 'backpropPredictor'

1. Input → a1, a2, a3, theta1, theta2, y, e, lambda
2. Output → [gradient vector]

- Reshape optimal theta vector into matrices for layer 1 and layer 2
- Apply forward propagation to make final predictions.
    1. input →x, optTheta1, optTtheta2

10) Alpha is specified (0.75) → learning rate
11) Classifier is called to train neural network

1 It takes input: array 'X', classification of outputs, alpha, IEPSILON.

2 It specifies some variables

- THETA1 = 0

- THETA2 = 0

- THETA3 = 0

- FC = 0

- J = 0

3 A function "nn" is called.

- It takes (X, Y, THETA1,THETA2, THETA3,IEPSILON,initw(initial state), alpha, J, FC) as input

- Thetas are calculated through following formula: -

- THETA1 = 2*IEPSILON* rand (K, number of features) – IEPSILON

- THETA2 = 2*IEPSILON* rand (K, K+1) – IEPSILON

- THETA3 = 2*IEPSILON* rand (K, K+1) – IEPSILON

- In a loop forward propagation is applied and if it is the '500' iteration then some calculations are done.

- 1. This forward propagation takes input X, THETA1 THETA2, THETA3.

- Then in the same loop back propagation is done through function "backpropclassification"

    1. Input→Y, A3,A2, α, N THETA3,THETA2.
    2. Output → DELTA1, DELTA2, DELTA3.

- Then in the same loop new theta values are calculated: -
  THETA1_NEW = THETA1-(alpha*DELTA1)
  THETA2_NEW = THETA2 – (alpha*DELTA2)
  THETA3_NEW = THETA3 – (alpha*DELTA3)

12) Test month → Testing results to make future predictions about weather features
13) Save yearly predicted data into text files.
14) Plot graphs for weather features of 20 years.

IV. RESULTS AND DISCUSSION

After designing and training the multi-layered neural network with the existing dataset, we have obtained a relationship between the existing non-linear parameters of weather. The result reveals that by increasing the number of hidden layers, the trained neural network can classify and predict the weather variables with less error. This could even be possible to predict aforementioned weather parameters on

monthly bases, and daily bases. Due to the increasing number of figures in a single study, we have focused only on the classification of weather parameters and yearly based prediction of weather parameters. We have obtained the analytical results through above described methodology and these results are explained with detailed description along with each diagram. Each diagram illustrates the actual and predicted results for one year. Due to the increasing number of hidden layers, trained neural network can classify and predict the weather variables with less error.

Fig. 3 graph, shows actual weather (sunny, foggy, thunderstorm, and rainy) and weather specified after classification by neural network over the course of 20 years (7000 days). Fig. 4 graph, shows actual and predicted mean wind speed over the course of 20 years (7000 days).
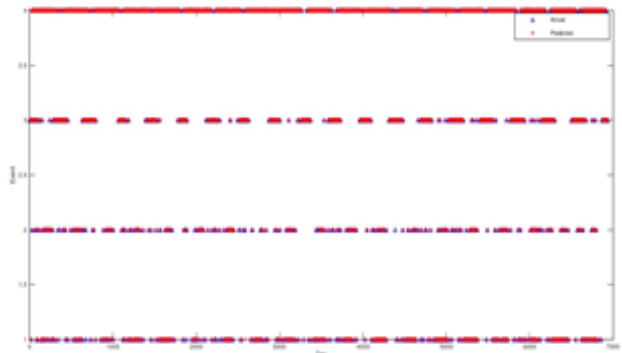

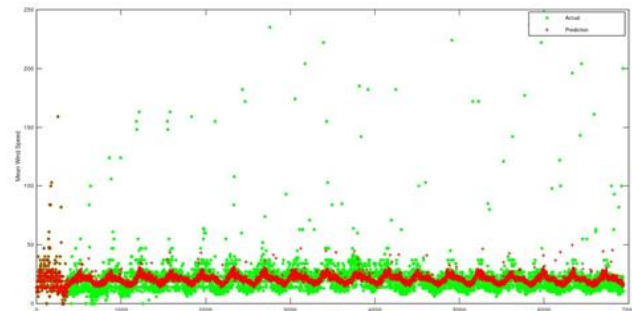Fig. 3.   Weather Event Classification.

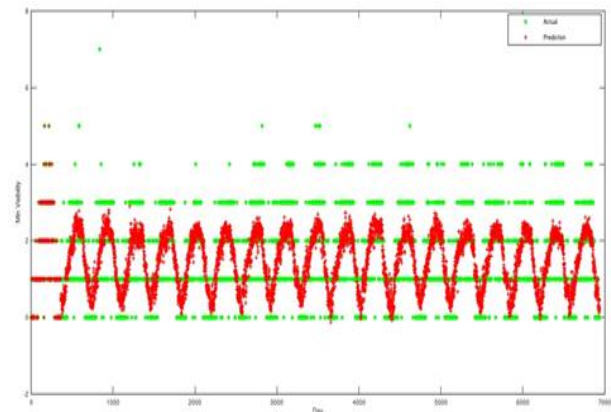Fig. 4.   Actual and Predicted Minimum Wind Speed.

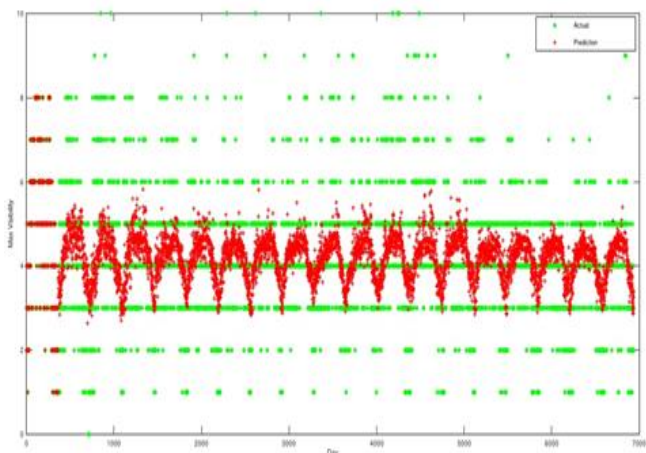

Fig. 5.   Actual and Predicted Minimum Visibility.

Fig. 6.   Actual and Predicted Maximum Visibility.

Fig. 5 graph shows actual and predicted min visibility over the course of 20 years (7000 days). Fig. 6 graph that shows actual and predicted max visibility over the course of 20 years (7000 days).

Fig. 7 graph that shows actual and predicted min pressure over the course of 20 years (7000 days). Fig. 8 graph that shows actual and predicted max pressure over the course of 20 years (7000 days).
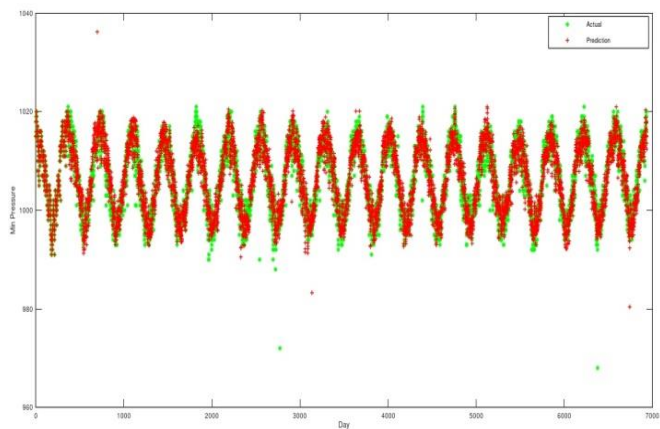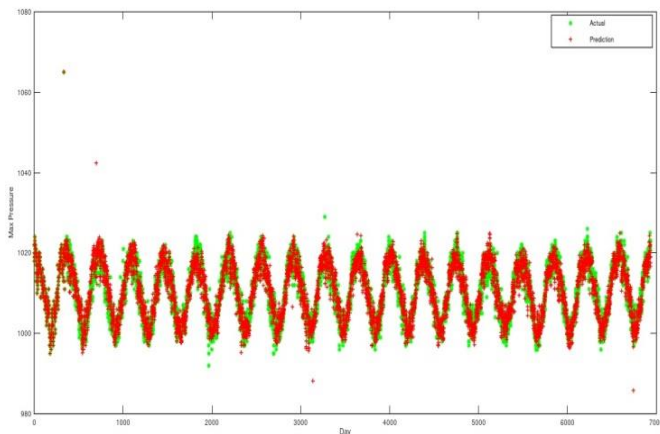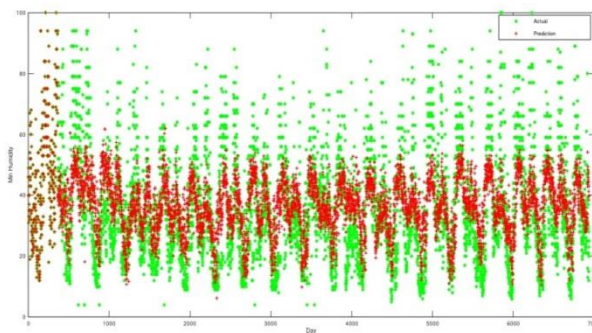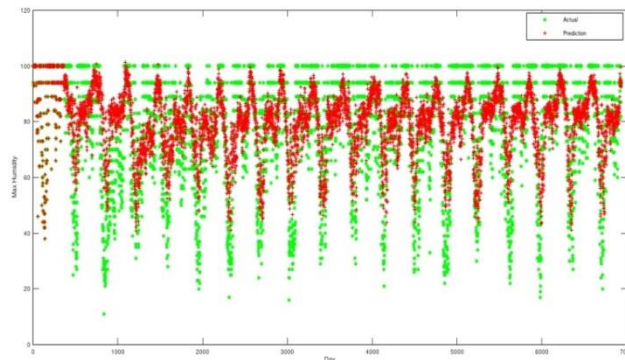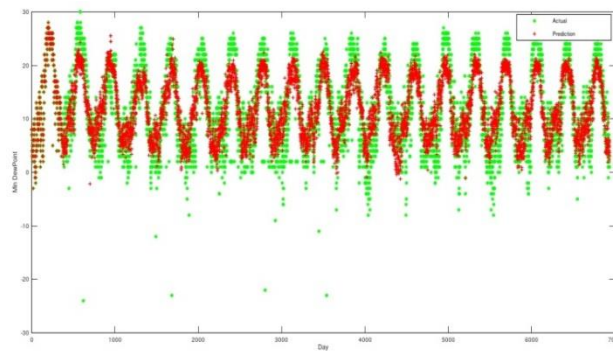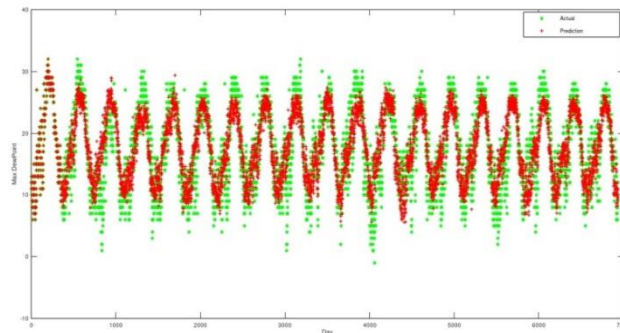

Fig. 7.   Actual and Predicted Minimum Pressure.

Fig. 8.   Actual and Predicted Maximum Pressure.

Fig. 9.   Actual and Predicted Minimum Humidity.

Fig. 10.   Actual and Predicted Maximum Humidity.

Fig. 9 graph shows actual and predicted min humidity over the course of 20 years (7000 days). Fig. 10 graph shows actual and predicted max humidity over the course of 20 years (7000 days).

Fig. 11.   Actual and Predicted Minimum Dewpoint.

Fig. 12.   Actual and Predicted Maximum Dewpoint.

Fig. 11 graph that shows actual and predicted min dewpoint over the course of 20 years (7000 days). Fig. 12 graph that shows actual and predicted max dewpoint over the course of 20 years (7000 days).

Fig. 13 graph that shows actual and predicted min temperature over the course of 20 years (7000 days). Fig. 14 graph that shows actual and predicted max temperature over the course of 20 years (7000 days).
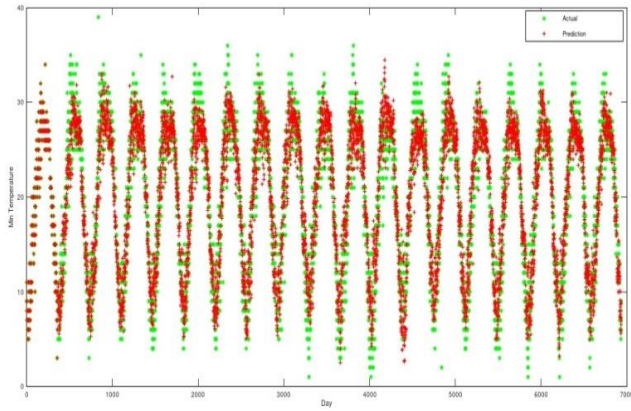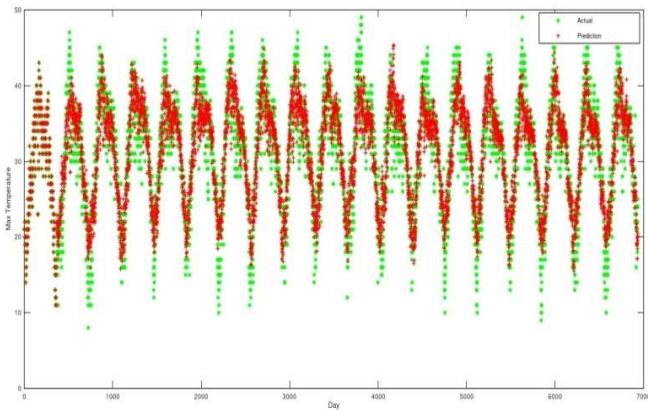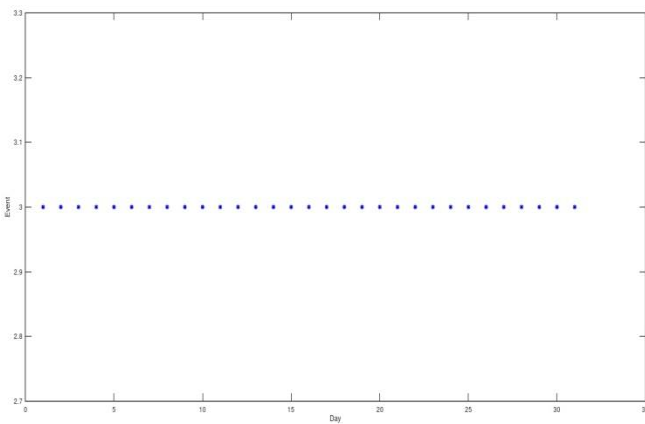

Fig. 13. Actual and Predicted Minimum Temperature.

Fig. 14. Actual and Predicted Maximum Temperature.

Fig. 15. Weather event Classification for One Month.
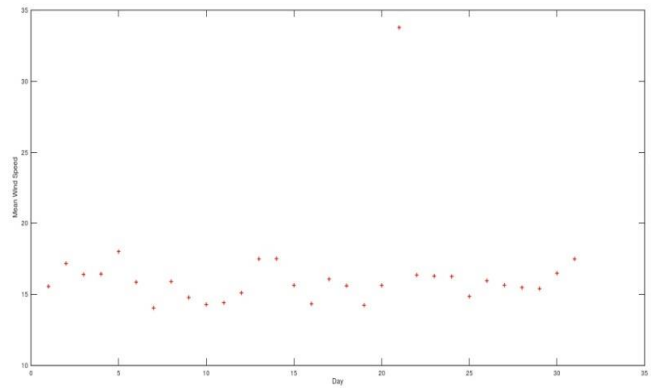
Fig. 16. Mean Wind Speed in Month Jan 2016.

Fig. 15 graph of classification of weather into rainy, thunderstorm, foggy, sunny (here these categories are converted to numbers and then categorized. The classification is done through the neural network which is trained by 20 years of data) in year 2016 over the course of 365 days. Fig. 16 graph of prediction of mean wind speed in year 2016 over the course of 365 days.

Fig. 17 graph of prediction of min visibility in year 2016 over the course of 365 days. Fig. 18 graph of prediction of max visibility in year 2016 over the course of 365 days.
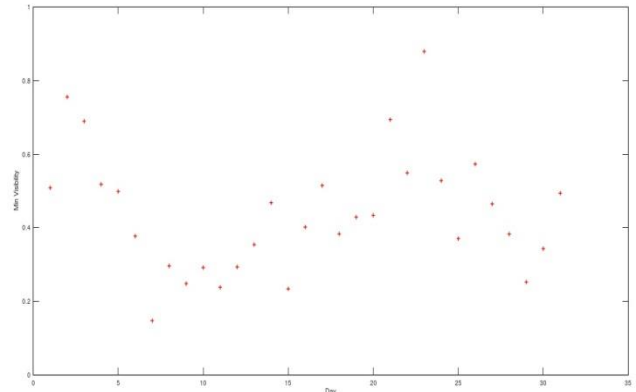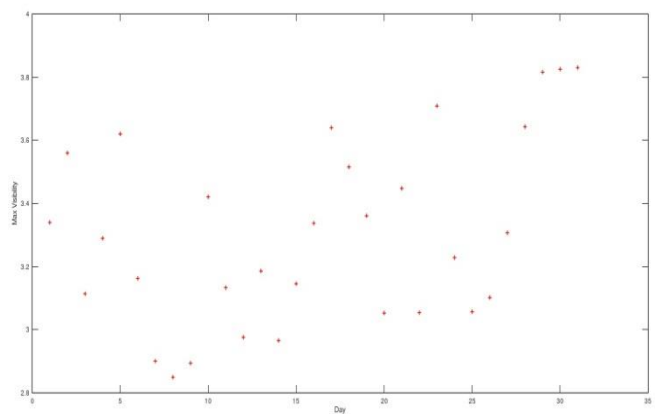

Fig. 17. Min Visibility in Month Jan 2016.

Fig. 18. Max Visibility in Month Jan 2016.

Fig. 21 graph of prediction of min humidity in year 2016 over the course of 365 days. Fig. 22 graph of prediction of max humidity in year 2016 over the course of 365 days.
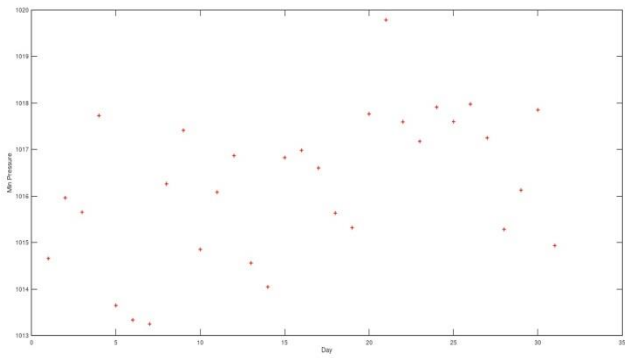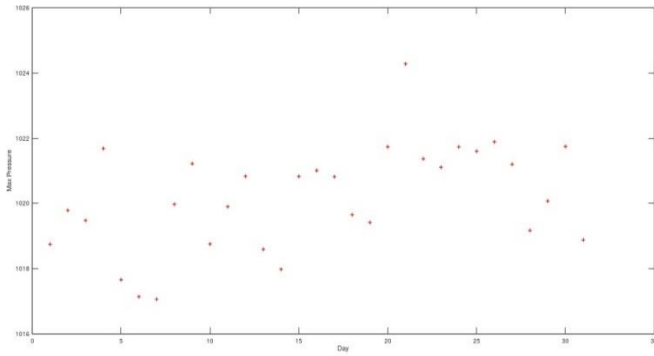

Fig. 19. Min Pressure in Month Jan 2016.

Fig. 20. Max Pressure in Month Jan 2016.

Fig. 19 graph of prediction of min pressure in year 2016 over the course of 365 days. Fig. 20 graph of prediction of max pressure in year 2016 over the course of 365 days.
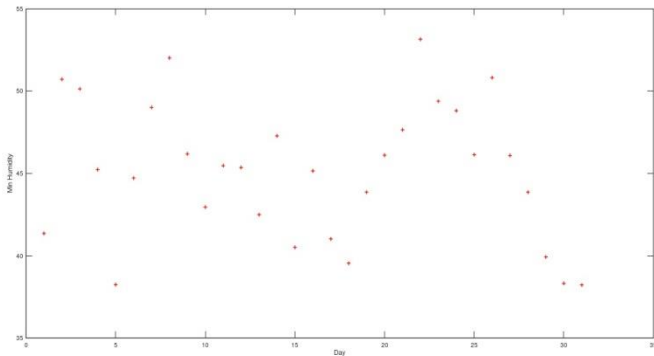

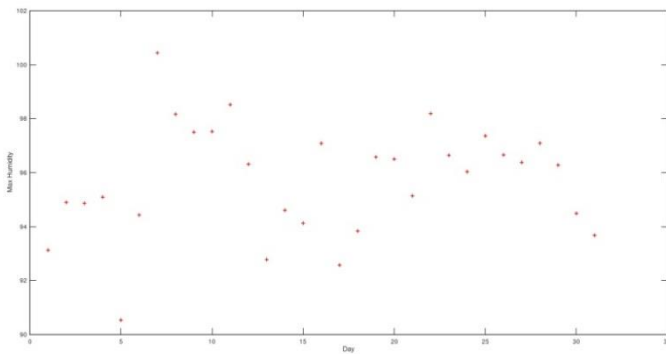Fig. 21. Min Humidity in Month Jan 2016.

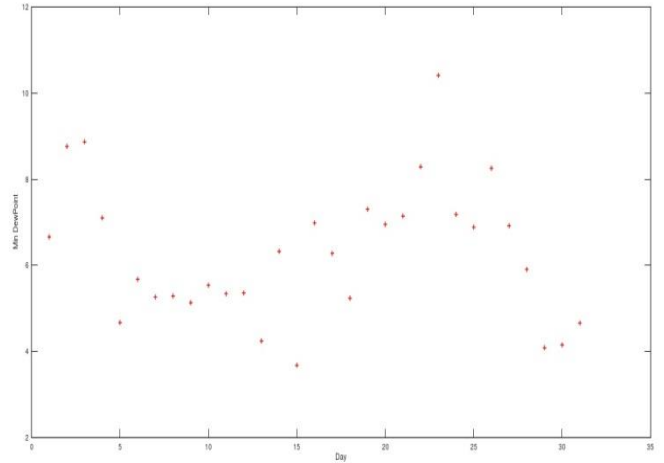Fig. 22. Max Humidity in Month Jan 2016.

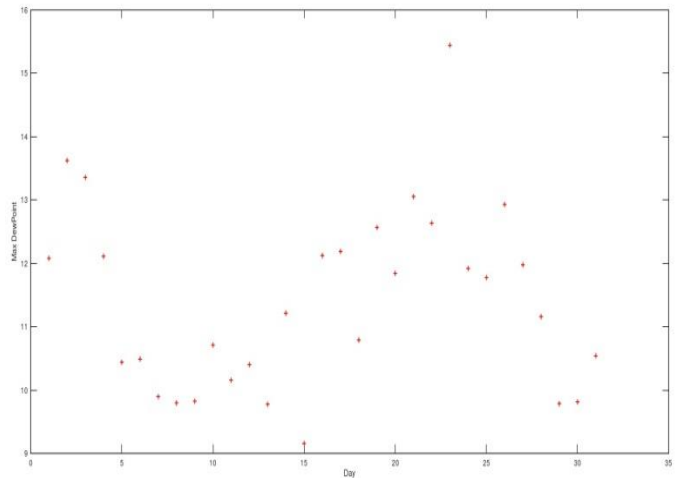Fig. 23. Min DewPoint in Month Jan 2016.

Fig. 24. Max DwePoint in Month Jan 2016.

Fig. 23 graph of prediction of min dewpoint in year 2016 over the course of 365 days. Fig. 24 graph of prediction of max Dewpoint in year 2016 over the course of 365 days.
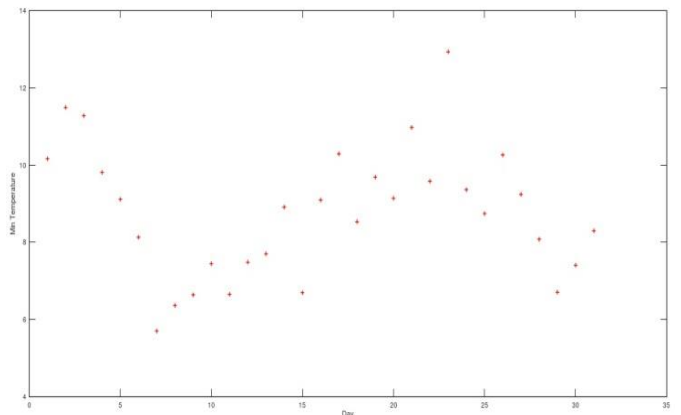

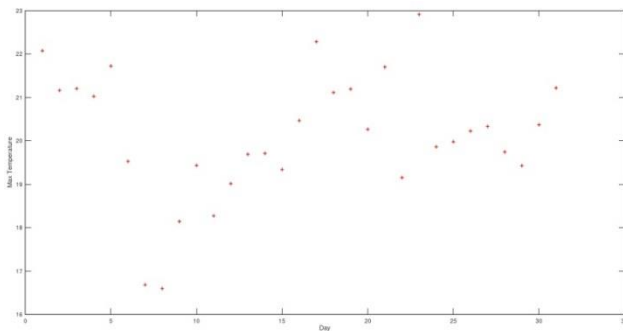Fig. 25. Min Temperature in Month Jan 2016.

Fig. 26. Max Temperature in Month Jan 2016.

Fig. 25 graph of prediction of min temperature in year 2016 over the course of 365 days. Fig. 26 graph of prediction of max temperature in year 2016 over the course of 365 days.

V. CONCLUSION

In this paper, a most powerful prediction algorithm called back propagation algorithm was used to predict and classify the weather forecast standard dataset. Eleven weather features were used to perform classification of weather into four types. Furthermore, twenty training examples from 1997-2015 were used to predict eleven weather features. The prediction was calculated for weather forecast basic factors like humidity, speed, etc. A Multi-layered neural network is designed and trained with the existing dataset and obtained a relationship between the existing non-linear parameters of weather. The overall behavior of our model has been concluded is that by increasing the number of hidden layers, the trained neural network can classify and predict the weather variables with less error. In upcommig paper, we will try to predict aforementioned weather parameters on monthly bases, and daily bases. Due to increasing number of figures in a single study, we have focused only on two objectives, i.e, classification and year based prediction. Additionally, a comparative analysis of several ANN techniques will be made so that we can emphasis our results broadly.

REFERENCES

[1] Delle Monache, Luca, Thomas Nipen, Yubao Liu, Gregory Roux, and Roland Stull. "Kalman filter and analog schemes to postprocess numerical weather predictions." Monthly Weather Review 139, no. 11 (2011): 3554-3570.

[2] Delle Monache, Luca, F. Anthony Eckel, Daran L. Rife, Badrinath Nagarajan, and Keith Searight. "Probabilistic weather prediction with an analog ensemble." Monthly Weather Review 141, no. 10 (2013): 3498-3516.

[3] De Giorgi, Maria Grazia, Antonio Ficarella, and Marco Tarantino. "Assessment of the benefits of numerical weather predictions in wind power forecasting based on statistical methods." Energy 36, no. 7 (2011): 3968-3978.

[4] Malik, Pooja, Saranjeet Singh, and Binni Arora. "An effective weather forecasting using neural network." Int J Emerg Eng Res Technol 2, no. 2 (2014): 209-212.

[5] Olaiya, Folorunsho, and Adesesan Barnabas Adeyemo. "Application of data mining techniques in weather prediction and climate change studies." International Journal of Information Engineering and Electronic Business 4, no. 1 (2012): 51.

[6] de Lima, Glauston R. Teixeira, and Stephan Stephany. "A new classification approach for detecting severe weather patterns." Computers & geosciences 57 (2013): 158-165.

[7] Sun, Xiaorong, Peter B. Luh, Kwok W. Cheung, Wei Guan, Laurent D. Michel, S. S. Venkata, and Melanie T. Miller. "An efficient approach to short-term load forecasting at the distribution level." IEEE Transactions on Power Systems 31, no. 4 (2016): 2526-2537.

[8] Anandharajan, T. R. V., G. Abhishek Hariharan, K. K. Vignajeth, and R. Jijendiran. "Weather Monitoring Using Artificial Intelligence." In Computational Intelligence and Networks (CINE), 2016 2nd International Conference on, pp. 106-111. IEEE, 2016.

[9] Araújo, Ricardo de A., Adriano LI Oliveira, Sergio Soares, and Silvio Meira. "Dilation-erosion perceptrons with evolutionary learning for weather forecasting." In Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on, pp. 3070-3077. IEEE, 2011.

[10] Islam, Taohidul, Sajal Saha, Ali Ahmed Evan, Nabonita Halder, and Shakti Chandra Dey. "Monthly Weather Forecasting through ANN Model: A Case Study in Barisal, Bangladesh." International Journal of Advanced Research in Computer and Communication Engineering 5, no. 6 (2016).

[11] Chen, Changsong, Shanxu Duan, Tao Cai, and Bangyin Liu. "Online 24-h solar power forecasting based on weather type classification using artificial neural network." Solar Energy 85, no. 11 (2011): 2856-2870.

[12] Purnomo, H. D., K. D. Hartomo, and S. Y. J. Prasetyo. "Artificial neural network for monthly rainfall rate prediction." In IOP Conference Series: Materials Science and Engineering, vol. 180, no. 1, p. 012057. IOP Publishing, 2017.

[13] Detwiler, Luke. "Using Backpropagation Neural Networks for the Prediction of Residual Shear Strength of Cohesive Soils." (2014).

[14] Narvekar, Meera, and Priyanca Fargose. "Daily weather forecasting using artificial neural network." International Journal of Computer Applications 121, no. 22 (2015).

[15] Balsamo, Gianpaolo, Rui Salgado, Emanuel Dutra, S. Boussetta, T. Stockdale, and Miguel Potes. "On the contribution of lakes in predicting near-surface temperature in a global weather forecasting model." Tellus A: Dynamic Meteorology and Oceanography 64, no. 1 (2012): 15829.

[16] Viswambari, M., and Dr R. Anbu Selvi. "Data Mining Techniques to Predict Weather: A Survey." International Journal of Innovative Science, Engineering & Technology 1, no. 4 (2014): 3.

[17] Sukanya, R., and K. Prabha. "Comparative Analysis for Prediction of Rainfall using Data Mining Techniques with Artificial Neural Network." Volume-5, Issue-6, Page 288-292 (2017).

[18] Maleki, Akbar, Morteza Gholipour Khajeh, and Marc A. Rosen. "Weather forecasting for optimization of a hybrid solar-wind–powered reverse osmosis water desalination system using a novel optimizer approach." Energy 114 (2016): 1120-1134.