© 2020 IJSRST | Volume 5 | Issue 7 | Print ISSN : 2395-6011 | Online ISSN: 2395-602X

4th National Conference on Advances in Engineering and Applied Science
Organized by : Anjuman College of Engineering and Technology (ACET) Nagpur, Maharashtra, India, In association with
International Journal of Scientific Research in Science and Technology

# Weather Forecasting System

**Ashley Mohanani, Rucha Kalode, Aishwarya Dhumane, Aman Zade, Shweta Hiwanj, Chanchla Tripathi**
Information Technology, DMIETR, RTMNU, Wardha, Maharashtra, India

## ABSTRACT

The Meteorological Department is an agency that takes readings of meteorological observations, weather forecasting and seismology. Universally, there are three types of weather prediction techniques, namely, synoptic weather prediction, numerical weather prediction, statistical weather prediction. The presented paper focuses on statistical weather prediction using the concept of artificial neural networks and regression algorithms. Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), Linear Regression and Random Forest Regression, are the algorithms that are used in the study, where the first three are categorized as ANN algorithms and remaining two are regression algorithms that do not include neural networks. A comparison is shown between all the five algorithms and their predicted value of relative humidity, with date and time as input values. This paper is used as a study of which algorithm provides best accuracy with the limited time and resource constraints.

**Keywords :** Artificial Neural Network, Recurrent Neural Network, Long Short Term Memory, Gated Recurrent Unit, Linear Regression, Random Forest Regression, Deep Neural Networks.

## I. INTRODUCTION

Weather Forecasting is an important application in meteorology department that predicts the weather conditions in the specified date and hour. In the last century, prediction of weather condition has been a very difficult task. To make an accurate prediction is one of the major challenges faced by meteorologist all over the world. From ancient times scientists have been trying different methods for predicting weather conditions some of those methods have been more accurate compared to others.

Weather forecasting predicts that how a present atmosphere will change. Present weather conditions are obtained by ground observations, observations from ships and aircraft, Doppler radar, and satellites. This information is sent to meteorological centres where the data is collected, analysed, and made into a variety of charts, maps, and graphs.

There are several weather prediction techniques such as numerical weather prediction, statistical weather prediction and. Meteorologists predicts the changes in weather patterns by using several tools such as radar, satellite and surface maps that measure temperature, air pressure, wind speed, and wind direction. This is a cycle that trains computer for weather forecasting.

We have used different techniques as Neural Network, it takes input and processes it using its hidden layers and produces the output. The Neural Network algorithms used are Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU). Also regression algorithms used are Linear Regression, and Random Forest Regression. We have analysed through all the

techniques and found out the best results that gives the best accuracy.

## II. Methods and Material

### A.Methodology

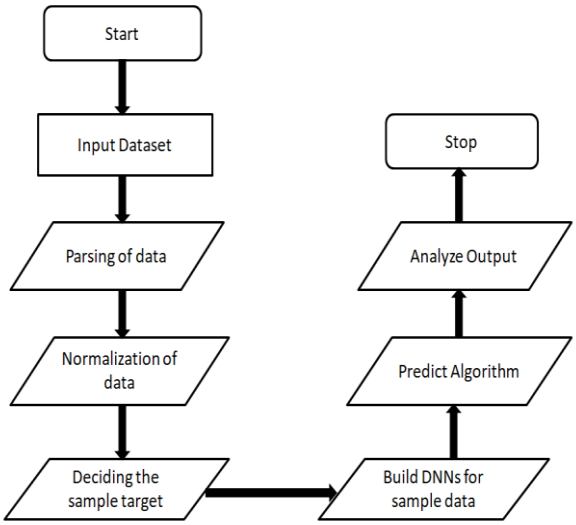


**Fig 1.** Methodology of the system

**i) Input dataset:**

Model to be trained, it needs to have some prior knowledge about the changes in the weather. The dataset has to be decided as per the training requirement of the neural network. Hence a large amount of dataset has to be given to the system on which it would be trained, validated and tested. Data will be processed to form the cycle to give accuracy of the output.

**ii)Parsing of data:**

Parsing the data means pre-processing the data. The raw data that was taken as an input can be a messed-up data, to separate out the data in particular category parsing is done so that the output data is a cleaner and simplified version of the input data. Parsing can be done at simple as well as large scale.

**iii)Build DNNs for sample data:**

The artificial neural network is derived from the biological neural network. A neural network recognises patterns, that includes minimum three layers i.e. input layer, processing layer and output layer. It trains itself by using these layers to give the accuracy in output.

Similarly, deep neural network is an advanced version of neural network. It has more depth because it has more hidden layers for processing.

**iv)Normalization of the data:**

The machine learning technique often applied as a part of data preparation is called as Normalization. The goal of normalization is to change the values of numeric columns in the dataset to a common scale, without distorting differences in the ranges of values. For machine learning, every dataset does not require normalization. It is required only when features have different ranges.

**v)Deciding the sample target:**

After pre-processing has been done, the data will be ready to be analysed. The data will be analysed on the basis of which weather parameter can be used; various parameters of weather can be max temperature, min temperature, humidity, rainfall etc. These are some parameters that are mostly worked on.

Deciding the parameter play a vital role in later on designing the network and predicting the algorithm. When the dataset has been analysed and parameters are decided, the next step will be to build the DNN.

**vi)Predict Algorithm:**

Data mining consist of various prediction algorithms used to predict someof future conditions. These algorithms are classified into clustering (Simple k

Means, Mean Shift, DBSCAN, etc.), And classifying (Logistic Regression, Naïve Bayes, KNN, etc.) and decision tree algorithms. (ID3, CART, C5, C4.5, etc.)

For these systems, algorithm will be decided based on the decided parameters and built neural network. Various combinations of algorithms will be tested and the algorithm with highest accuracy will be selected. The accuracy can be calculated by comparing the predicted weather condition with the actual weather.

The various prediction algorithm used for the weather forecasting is

- Recurrent Neural Network.
- Long Short-Term Memory.
- Gated Recurrent Units
- Random Forest Regression.
- Linear Regression.

**vii)Analyse Output:**

After building, training and testing have been done, the network will be ready to generate output. But the output has to be accurate for the system to be useful. Hence the accuracy has to be verified by comparing the outputs with other sources.

## B. PREDICTION ALGORITHMS

Prediction algorithms are those that helps in predicting the output in future using algorithms. Some of the algorithms used are –

**i)Recurrent Neural Network**:

RNN is a neural network that takes input as raw data processes it in the hidden layers and gives output. Then obtained output is compared to the expected output and the difference is found out, called loss. And then the cycle repeats until we obtain the least loss.

**ii)Long – Short Term Memory:**

LSTM is a Neural Network that can process single data as well as sequence of data. It is composed of a cell, an input gate, and a forget gate. LSTM network is best suited for classifying, processing and making predictions based on time series data.

$$\tilde{c_t} = tanh(w_c[h_{t-1}, x_t] + b_c)$$
$$c_t = f_t * c_{t-1} + i_t * \tilde{c_t}$$
$$h_t = o_t * tanh(c^t)$$

**iii)Gated Recurrent Unit:**

GRU is a recurrent neural network that use connections through a sequence of nodes to perform machine learning tasks. It overcomes the problem of vanishing gradient. It has two gates update gate and reset gate that are used for controlling the flow of information over a period of time.

**iv)Linear Regressor:**

Linear Regressor is a data analysis technique that is used to determine the relationship between a dependent variable and independent variable using a straight line. It is useful while prediction or forecasting or reduction. It uses regression equation that is used for finding out relationship between sets of data.

**v)Random Forest Regressor:**

Random Forest Regressor is a meta estimator that built multiple decision trees and merge their predictions together to get more accurate and stable output. While training each tree gets trained from random samples of training observations.

## C. Material

The Dataset istaken from city Jena, of Germany. The duration of the dataset is from 1st of January 2009 to

31st December 2016 at the time interval of the dataset to working of weather forecasting is 10 min.

## III. Result and Discussion

In this paper, results are shown on how various algorithms have worked after training, testing and predicting. The result section contains line graph which shows the difference between expected and predicted value. Another graph shows the loss value decreasing during the training period and a graph that shows final prediction values of relative humidity.

### 1. Long Short Term Memory

a. Line Graph

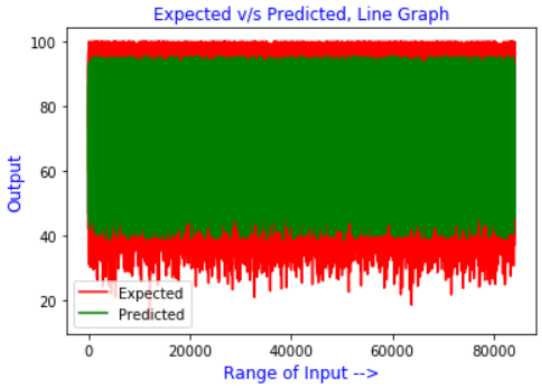


Fig. LSTM Expected v/s Predicted Line Graph

The graph shows expected values in red and predicted values in green colour line with x axis as range of input and y axis as output. The range of predicted value expands from 35 to 95 whereas the expected value ranges from 20 to 100.

b. Loss Graph

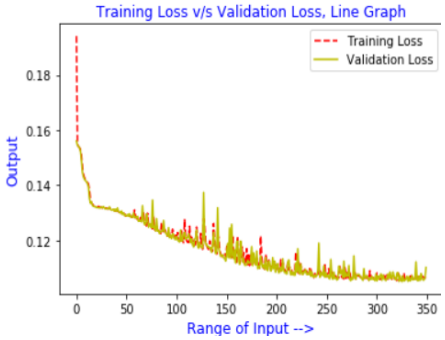


Fig. Loss Graph

The graph shows the downfall of loss function while the model was in its training process. Loss function used in the LSTM model is 'mean absolute error'.

c. Humidity Graph

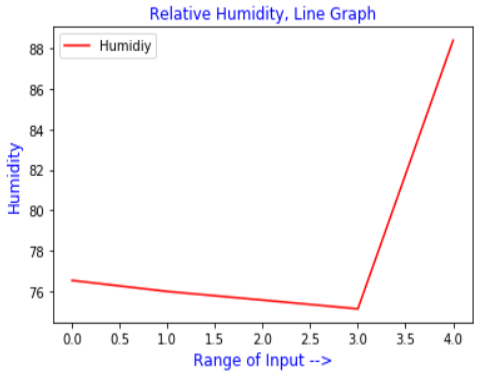


Fig. Humidity Predicted Graph

The final graph shows the variation in the predicted values of relative humidity. These values are of same day with the difference of timestamp.

d. Predicted Values

| | rh (%) |
|---|---|
| 0 | 76.528214 |
| 1 | 75.982452 |
| 2 | 75.556137 |
| 3 | 75.120827 |
| 4 | 88.416725 |

Fig. Predicted Values by LSTM

Above table shows the predicted values of relative humidity given after training of LSTM algorithm.

2. Recurrent Neural Network

a. Line Graph

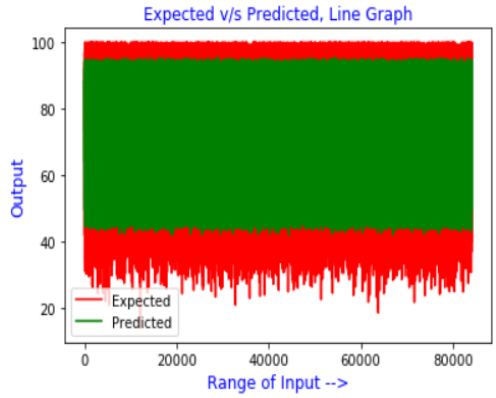


Fig.RNN Expected v/s Predicted Line Graph

The line graph shows various values when passed through the neural network and results produced by it v/s the expected results. As per the results, the lower values are a little less accurate as compared to the upper values.

b. Loss Graph

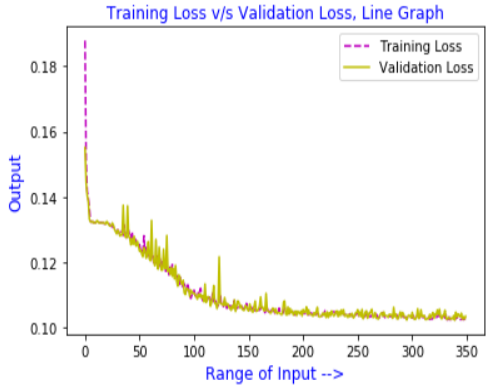


Fig. Loss Graph

The decreasing loss graph of RNN shows fluctuations as the model gets more accurate. Training and testing loss has been plotted in the above graph.

c. Humidity Graph

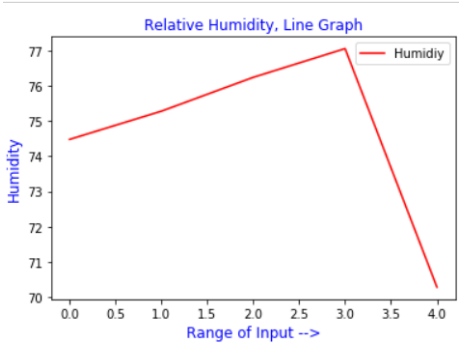


Fig. Humidity Predicted Graph

The humidity graph shows the predicted values of humidity for RNN network.

d. Predicted Values

| | rh (%) |
|---|---|
| 0 | 74.477776 |
| 1 | 75.280373 |
| 2 | 76.242119 |
| 3 | 77.058464 |
| 4 | 70.277946 |

The table above contains the final predicted values of humidity by RNN network.

## 3. Gated Recurrent Unit

a. Line Graph

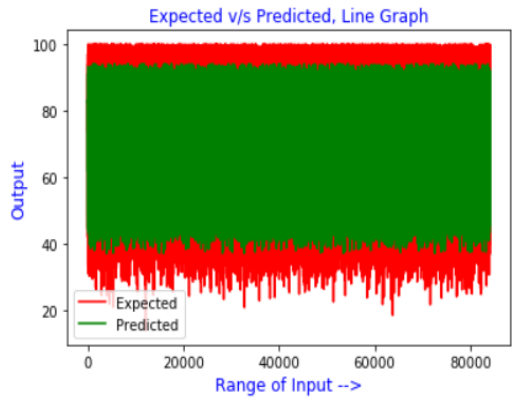


Fig. GRU Line Graph

GRU line graph shows the difference between predicted and expected output where the range of expected data is a little larger than that of predicted data.

b. Loss Graph

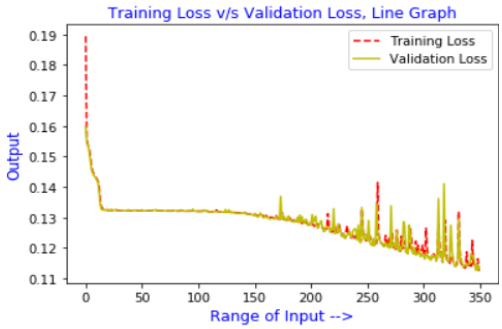


Fig. Loss Graph

The loss graph shows more fluctuation in terms of validation loss, which is shown by yellow solid line whereas the training data is shown in red dashed line.

c. Humidity Graph

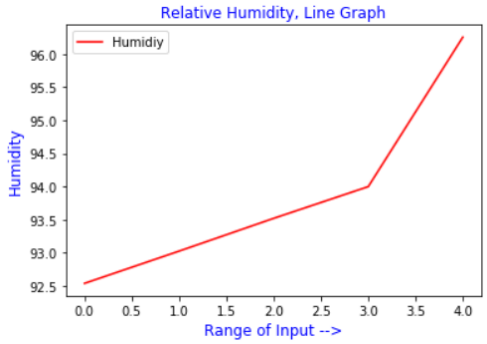


Fig. Humidity Graph

The final humidity graph shows predicted values of relative humidity by GRU algorithm. Final values ranges from 92 to 96.

d. Predicted Values

| | rh (%) |
|---|---|
| 0 | 92.537308 |
| 1 | 93.024559 |
| 2 | 93.519508 |
| 3 | 93.998741 |
| 4 | 96.257019 |

Fig. Predicted Values

The table above shows the final predicted values of relative humidity by GRU at different timestamps.

## 4. Random Forest Regression

a. Scatter Graph

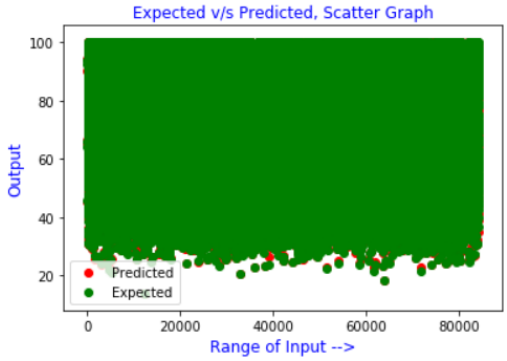


Fig. Random Forest Scatter Graph

The given scatter graph shows red dots as predicted values and green dots as expected values.

b. Humidity Graph

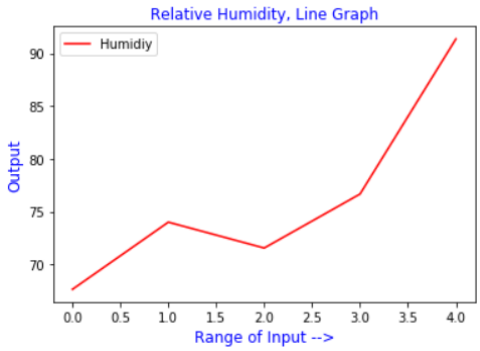


Fig. Humidity Graph

The final graph above depicts predicted values of humidity as per random forest regression algorithm.

c. Predicted Values

| | rh (%) |
|---|---|
| 0 | 67.652467 |
| 1 | 74.009333 |
| 2 | 71.561667 |
| 3 | 76.672767 |
| 4 | 91.354333 |

Fig. Predicted Values

The table shows the values predicted by Random Forest Regression algorithm.

**5. Linear Regression**

a. Line Graph

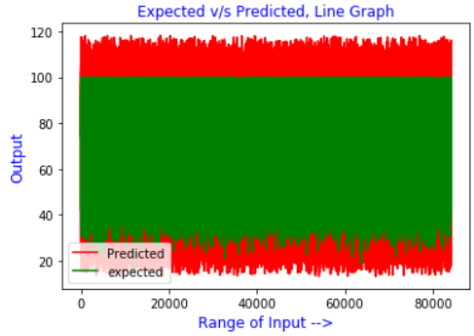


Fig.LSTM Expected v/s Predicted Line Graph

The graph shows expected values in red and predicted values in green colour line with x axis as range of input and y axis as output.

b. Humidity Graph

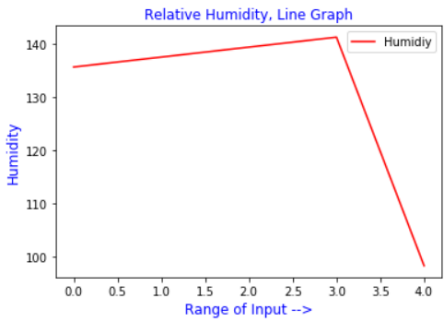


Fig. Humidity Graph

The final graph shows the predicted values of relative humidity which ranges from 98 to 135.

c. Predicted Values

| | rh (%) |
|---|---|
| 0 | 135.598584 |
| 1 | 137.468458 |
| 2 | 139.338333 |
| 3 | 141.208208 |
| 4 | 98.221987 |

Fig. Predicted Values

The values predicted by Linear Regression algorithm.

## IV. CONCLUSION

To conclude, below is the actual reading of relative humidity that was actually required while prediction.

| | Relative Humidity |
|---|---|
| 0 | 74 |
| 1 | 80 |
| 2 | 84 |
| 3 | 86 |
| 4 | 90 |

Fig. Expected Values

1. LSTM has proved to be the best working algorithm while prediction with a Mean Absolute Error (MAE) rate of 0.0958
2. The test graph is somewhat very close of LSTM and GRU, but here too LSTM graph has proven to be better than GRU, where GRU has an error rate of 0.1134
3. Random Forest Regression has shown a loss

rate of 0.099

4. Whereas, RNN has a loss rate of 0.112

LSTM has proven to be best for prediction of relative humidity after model has been properly trained and tested.

## V. REFERENCES


[1]. Adesesan . B Adeyemo, Olaiya Folorunsh "Application of Data Mining Techniques in Weather Prediction and Climate Change Studies" I.J. Information Engineering and Electronic Business, 2012, 1, 51-59 February 2012.

[2]. Aishwarya Dhore, Anagha Byakude, Bhagyashri Sonar, Mansi Waste "Weather prediction using the data mining Techniques" International Research Journal of Engineering and Technology (IRJET) Volume: 04 Issue: 05 May -2017.

[3]. M Ramzan Talib, Toseef Ullah, M Umer Sarwar, M Kashif Hanif and Nafees Ayub "Application of Data Mining Techniques in Weather Data Analysis" IJCSNS International Journal of Computer Science and Network Security, 22 VOL.17 No.6, June 2017.

[4]. Divya Chauhan, Jawahar Thakur "Data Mining Techniques for Weather Prediction: A Review" International Journal on Recent and Innovation Trends in Computing and Communication Volume: 2 Issue: 8 August 2014.

[5]. Vivienne Sze, Yu-Hsin Chen, Tien-Ju Yang, Joel Emer "Efficient Processing of Deep Neural Networks: A Tutorial and Survey" IEEE 13 August 2017.

[6]. Prashant Biradar, Sarfraz Ansari, Yashavant Paradkar, Savita Lohiya "Weather Prediction Using Data Mining International Journal of Engineering Development and Research Volume 5, Issue 2 2017.