JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

WEATHER FORECASTING

[1]Anupam Kumar Maurya, [2]Anurag Singh, [3] Er. Preeti Navel

[1]Undergraduate, [2] Undergraduate, [3]Assistant Professor

[1,2,3]CSE Department,

[1,2,3]SRMCEM, Lucknow, India

[1]anupammaurya264@gmail.com, [2]scanurag@gmail.com, [3]er.preetinaval09@gmail.com

Abstract: Weather forecasting has been playing an important factor due to its applications in the various sectors such as in agriculture, utility companies and in day-to-day life. Weather prediction is a real-time challenging issue witnessed by the world in the last decade. The prediction is becoming more complex due to the ever-changing weather conditions. Weather forecasting entails predicting how the present state of the atmosphere will change. For effective analysis of the weather, it is necessary to understand various influencing factors that cause the weather changes. Weather forecasting is the process of recording the parameters of weather like wind direction, wind speed, humidity, rainfall, temperature etc. Since machine learning techniques are more robust to perturbations, in this project we applied linear regression and Neural Network to predict the weather such as temperature, rainfall etc. and compare both approaches and analyze it.

Keywords— weather forecasting; Machine Learning(ML); Algorithms; Linear Regression; Neural Network

I. INTRODUCTION

Weather forecasting has become an important field of research in the last few decades. In most of the cases the researcher had attempted to establish a linear relationship between the input weather data and the corresponding target data. But with the discovery of nonlinearity in the nature of weather data, the focus has shifted towards the nonlinear prediction of the weather data. Although there are many literatures in nonlinear statistics for weather forecasting, most of them require that the nonlinear model be specified before the estimation is done. But since the weather data is nonlinear and follows a very irregular trend, Artificial Neural Network (ANN) has evolved out to be a better technique to bring out the structural relationship between the various entities. The paper examines the applicability of Artificial Neural Network approach by developing effective and reliable nonlinear predictive models for weather analysis. It also compares and evaluates the performance of the developed models using different transfer functions, hidden layers and neurons to forecast weather for certain time frame.

II. LITERATURE SURVEY

Over the last few decades, there has been a substantial amount of research done using numerical weather data. The data was primarily used for applying machine learning techniques. Other techniques included fuzzy logic and data mining. The majority of weather forecasting relies on generative approaches and the underlying principals are based on numerical methods.

Some people applied comprehensive tree-based learning algorithm. For example, N. Hasan and M.T Uddin used a tree-based algorithm, namely C4.5 and their output result had 96% accuracy. They also used Naïve Bayes, but the accuracy of C4.5 was much better in terms of f-score. [1] Lin and Chen [2] worked on a typhoon rainfall forecasting model using ANN and their result shows that excessive spatial rainfall information may not increase the generalization of the forecasting model. Awan and Awais [3] also tried to predict weather events based on a fuzzy RBS method for Lahore, Pakistan. They used two different datasets of 365 examples with only 4 features, and 2500 examples with 17 features. They mentioned their finding that fuzzy RBS method was sensitive to random sampling with replacement technique that was applied to produce. Another article reviews we found useful is S. B. Kotsiantis [4] mentioned few statistical classifiers to build classification tress. Using information entropy from a set of training examples of pre-classified samples, where each sample comprises of N-dimensional vector. H. Zhang, X. Zhao and S.Zou [5] proposed a classification algorithm naming Neuron Classification Algorithm (NCA). The algorithm has higher approximation function. They introduced the law of attraction here which increases the accuracy of weather forecast. It has been known to classify the test samples more accurately than Euclidean distance. As dataset, they chose forecast of abnormal megathermal weather in North of Zhejiang province. Combining Neural Networks and ARIMA Models for Hourly Temperature Forecast H. S. Hippert , C. E. Pedreira and R. C. Souza in their works [7] on hourly temperature forecasting, proposed to use the combination of Neural Networks and ARIMA models. The forecast is done on the basis of previous temperature records, maximum and minimum temperature data supplied by weather service. The previous temperatures are used as input to the Artificial Neural Network (ANN) where there is only one output node, the predicted temperature at a particular time. Their results show that hybrid systems based on ARMA models produce more accuracy than auto progressive models.

Simple classifiers such as Support Vector Machines (SVM) or Artificial Neural Networks (ANN) are also widely used to classify certain parameters such as rainfall and cloud states.

III. DESCRIPTION OF DATASET

The dataset was obtained from KAGGLE. The data set is comprised of the following variables:

Temperature(ºC): The temperature variable has a great impact on precipitation and is greatly related to humidity (%). The dataset has both maximum and minimum temperatures along with the mean values.

Cloud (okta): The cloud variable measured in Okta impacts the Earth's surface by reflecting incoming sunlight. It is also responsible for absorbing the heat emitted from the surface and radiating to space. The dataset comprises daily cloud data measured in a range from 0-8 Okta.

Wind Speed (knots): The wind variable measured in knots shows how quick the air is moving. The wind speed also has a direction and has various impacts on surface water and evaporation. The dataset comprises daily prevailing wind speed.

Rainfall (mm): The rainfall variable is a very important metric in weather forecasting. It helps the environment to continue to stay in its position the way it should be. Agriculture of Bangladesh mostly depends on rainfall. The dataset has a daily basis of total rainfall data in millimetres.

Sunshine (hour): The sunshine variable measures the amount of sunshine at a particular place. The Sun is the basic cause of our changing weather. The day-night cycles in the weather have obvious causes and effects on weather.

Humidity (%): The humidity variable measured in percentage helps to calculate the amount of moisture in the air at a given time on a given day, which is simply the ratio of water vapor and dry air.

Sea Level Pressure (millibars): The sea level pressure variable plays a significant role in the formation of weather conditions in a certain area. It is a component that has mass and weight. This means a vast ocean of air inserts a huge amount of pressure. So, it is natural that the air will affect the Earth's weather.

IV. OBJECTIVES

The objective of the research study are:

- To examine the applicability of Neural Network approach by developing effective and efficient predictive model for weather analysis
- To develop an efficient, reliable and effective weather forecasting system based on Neural Network and machine Learning.
- Study of the different techniques in machine learning that can be used to improve the prediction of weather.
- To compare and evaluate the performance of above models .

After observing about different methods of weather forecasting techniques we can understand techniques we can understand that by using we will get more accurate result. Then we will be also reduce the amount of error by which we can predict more accurately because the whole idea of weather forecasting is based on the accuracy of the machine learning model.

V. PROPOSED SYSTEM

Our model collects historical weather data that includes a variety of important factors that influence weather change, such as temperature, both maximum and minimum temperatures, atmospheric moisture or humidity, precipitation. In our proposed model, the collected dataset is divided into sections that are useful to the machine learning model and parts that aren't. After that, the dataset goes through data preprocessing, which involves passing the data through a process that replaces missing and error values in the dataset with mean values or the most frequently occurring value in that field. Following the data preprocessing, the cleaned dataset is divided into two parts: the training set and the test set. The training set is used to teach the machine learning model how to compute the results, while the testing set is used to find the results, compare the real and measured values, and use the error value as a benchmark to teach the machine learning model even more. The training process will also include fold cross validation, in which the dataset is divided into k sets k times and then into test and training sets, with training sets chosen at random in each set and the model trained in each set. The test set for a qualified machine learning model is then the kth set. This technique not only aids in the reduction of under fitting, but also the reduction of over fitting.

VI. PREDICTION ALGORITHMS

- We prefer to use Neural Networks algorithm to obtain accurate weather prediction.
- A neural network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates.
- A neural network will consist of three types of layers:
  1. The input Layer: It receives all the inputs and the last layer is the output layer which provides the desired output
  2. The Hidden Layer: All the layers in between these layers are called hidden layers. There can be a number of hidden layers. The hidden layers and perceptron in each layer will depend on the use-case you are trying to solve.
  3. The Output Layer: It provides the desired output. Upon which all the input is received and it forwarded to the next layer.
- We will be implementing Neural Networks algorithms on the available datasets and we will be also analyzing the prediction trends on different types of datasets. Most data sets have been obtained from Kaggle.
- Depending on the location of the weather station, it offers historical dataset as well as current dataset.
- The dataset given by the source is fairly accurate and include a variety of features to work with such as mean temperature, maximum temperature humidity and precipitation
- The dataset that we are working with is for a particular area and organized by date.

VII. RESULT EVALUATION

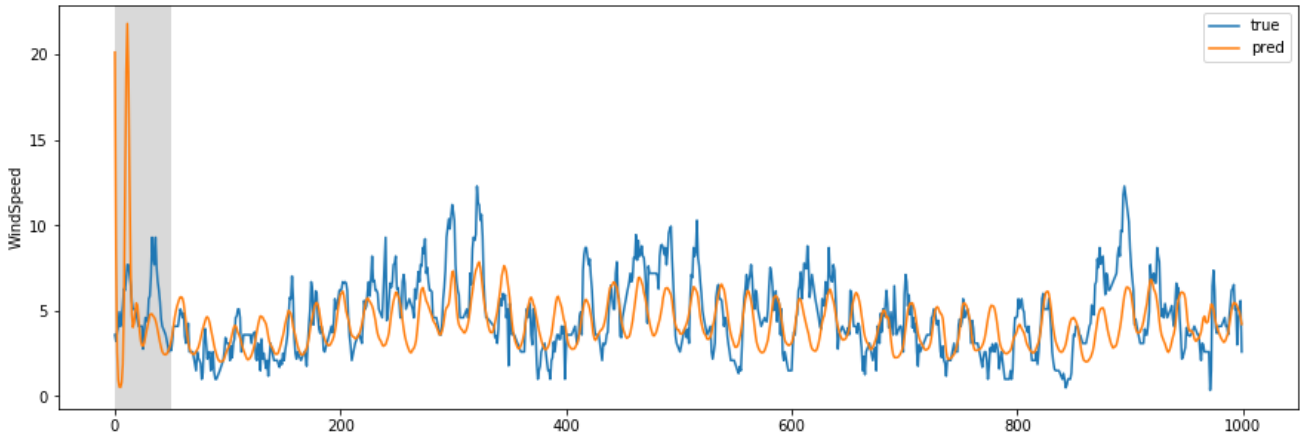

Fig1 Comparison between actual and predicted speed of the wind

This is comparison between the comparison between the component of the weather i.e wind speed.

The graph depicts the accuracy of the model.

Fig.1 describe the predicted and the actual Neural Network Model calculate the prediction with a good Accuracy.

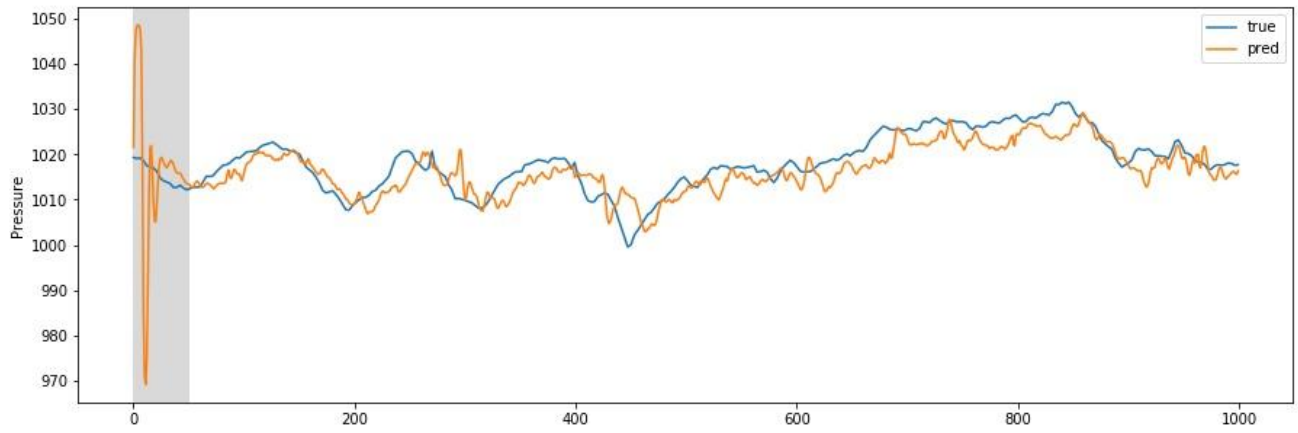

Fig2 Comparison between actual and predicted pressure of the atmosphere.

To validate the accuracy of the model We can also compare another weather parameter to ensure that model predict with a good accuracy

After plotting the graph,comparing between the predicted and actual result. We see that model works with wide range of the weather parameters.

This paper focuses on the trend of prediction technologies such as Machine Learning and Neural Network by training and testing of the previous values.Factors considered are Wind Speed and Pressure and can be also applied on Temperature and Humidity etc.

VIII. CONCLUSION

- Weather prediction is very challenging due to consistently changing of weather which are dependent on multiple parameters which form a complex pattern to identify
- The Historical dataset available on the website consists of all parameters such as temperature, pressure, rainfall etc. which is not fully sufficient.

- To obtain higher accuracy in the weather prediction we can also use new weather parameters rather than using existing weather parameters.
- Neural Network has been implemented in this project. The techniques have shown a better result in accuracy of the results. It has led to result that it is possible to predict the weather with more accuracy and efficiency using machine learning techniques. In the future
- Weather Forecasting has a big challenge of predicting the accurate results which are used in many real time systems like electricity departments, airports, tourism centers, etc. The difficulty of this forecasting is the complex nature of parameters. Each parameter has a different set of ranges of values. This issue is addressed by Neural Networks. It accepts all complex parameters as input and generates the intelligent patterns while training and it uses the same patterns to generate the forecast.

IX. ACKNOWLEDGMENT

We would like to thank all the working staff of Shri Ramswaroop Memorial Group of Professional College of Engineering and Management and a great thank to our project guide Er. Preeti Navel who gave her best to make us do this great project. Moreover, thanks to all the authors whose papers and books we referred during this project. Without help of all these resources, we would have never completed this project.

REFERENCES

[1] Y. Xu, V. Ramanathan, D. G. Victor, Global warming will happen faster than we think, Nature, Vol. 564, 30:32, 2018, Dec.
[2] D. Spratt, I. Dunlop, Existential climate-related security risk: A scenario approach, 2019 May, Policy Paper from Breakthrough-National Centre for Climate Restoration
[3] Aditya Grover, Ashish Kapoor, and Eric Horvitz.2015. A deep hybrid model for weather forecasting. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 379–386.
[4] Divya Chauhan, Jawahar Thakur "Techniques for Weather Prediction: A Review 5, India: ISSN, 2013.
[5] Jitcha Shivang, S.S Sridhar – Weather Prediction For Indian Location Using Machine Learning
[6] Kumar Abhishek, Abhay Kumar "A Rainfall Prediction Model Using Artificial Neural Network" department of Information Technology IEEE Control and System Graduate Research Colloquium (ICSGRC 2012),
[7] C.N. Schizas'BJ S. Michaelides' C.S. Pattichis3 R.R. Livesay', Artificial Neural Network in Forecasting MinimumTemparature,'University of Indianapolis, U.S.A
[8] M. Viswambari, Dr. R. AnbuSelvi, "Data Mining Techniques To Predict Weather: A Survey", ISSN 2348 – 7968, IJISETInternational Journal Of Innovative Science, Engineering & Technology, Vol. 1 Issue 4, June 2014.