International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 03 | Mar 2020 www.irjet.net p-ISSN: 2395-0072

Intelligent Weather Forecasting using Machine Learning Techniques

Prathap N Kashyap[1], Preetham.M.S[2], Rohan Venkatesh Nayak[3], Uday.B[4], Radhika.T.V[5]

[1,2,3,4]*Student, Dept. Of Information Science & Engineering, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India*

[5]*Assistant Professor, Dept. Of Information Science & Engineering, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India*

Abstract - *Weather is one of the complex processes to predict. Forecasting the weather is a highly challenging task. The purpose of this paper is to forecast the weather using machine learning techniques. In any machine learning technique, the most important thing for any model is data. With proper and clean data we can use many models to accurately predict the weather. We have proposed a weather forecasting techniques that uses the historical data of Bangalore City from Dark Sky website to train different machine learning models and can provide predictions about weather conditions.*

Key Words: **Weather forecast, Machine learning, Deep learning, regression, Neural network, RNN**.

1.INTRODUCTION

Weather forecasting can be defined as an attempt to predict the future weather conditions based on previously collected data. Sometimes extreme weather conditions can cause heavy losses. If we can predict the future weather conditions properly such losses can be minimized. Weather forecasting is important to each and everyone, whether it is a student who decides to carry an umbrella or not by knowing the weather, or a government organisation which helps people to vacate a location by knowing that it is going to rain heavily in that region. Existing systems give information about weather in terms of a wide ranged value. For example, the temperature is going to be ranging from 21 degrees to 29 degrees celsius. Such systems are very confusing and are not very helpful to the people.There are various fields which need weather prediction data. People who consume weather prediction results may include farmers, pilots, power generation stations which depend on solar energy and wind energy and many more.

There is a need for systems which can predict the weather conditions very accurately at a specific time and location. Machine Learning models can be used to build systems that can predict the weather with high accuracy.. The most important part of any machine learning models is the data. The prediction accuracy of a machine learning model is completely dependent on the data that it is trained with. Techniques of machine learning such as regression which take in the previous weather condition data are found to be pretty good at providing a decent level of accuracy at predicting the future results. One such regression techniques is multi target regression which has multiple dependent variables to be predicted. Neural Network approach for weather prediction also give promising results. Recurrent Neural Network is one such method which provides good results in predicting the future values by using the hidden layers and retraining the model until satisfying level of accuracy is reached by changing the weights given to the layers.

2. LITERATURE SURVEY

In the paper presented by A H M Jakaria et.al [1], a study was made by collecting the weather data of Nashville in Tennessee and data of surrounding cities. The training data set consisted of two months worth of weather data of July and August 2018. Many machine learning models were implemented such as Extra Tree Regression, Random Forest Regression, Support Vector Regression and Ridge Regression. They have found Random Forest Regressor to be a better regressor as it ensembles multiple decision trees while making decision. Their evaluation results have shown that machine learning models can give accurate results comparable to the traditional models.

In a study conducted by Man Galih Salman et al [2], a comparison was made among different deep learning models such as Convolutional Networks, Conditional Restricted Boltzmann Machine, Recurrent Neural Network. The dataset for training and testing the models have been collected from the indonesian Meteorological department. Rainfall is the feature being predicted while keeping the other variables independent. RNN method was found to give adequate accuracy when compared to the other models.

In a study conducted by Dires Negash Fente et al. [3], the LSTM algorithm has been used. Different combinations of weather parameters such as pressure,

temperature, dewpoints wind speed, precipitation and many other weather parameters were used to train the neural network model. The data required for training the model has been obtained from National Climate Data Center from the year 2007 to 2017. After using the data to train the model using LSTM algorithm, it was found that LSTM algorithm gave substantial results accuracy wise, among other weather prediction techniques.

In the paper presented by E B Abhrahamsen et al. [4] , two different neural network models were studied. One is an Artificial Neural Network and the other is an Artificial Neural Network with Exogenous input. Four different models were built each with a prediction period of 1, 3,6 and 12 hours. All the ANN models use the ReLu as the activation function in the hidden layer and a linear activation function in the output layer. IN the ANN models only temperature was used as the input to the network. Where as in ARX models another feature, precipitation was introduced as another input and improved the prediction accuracy. So they have concluded that with the introduction of more inputs accuracy of these models can be improved.

2.1 EXISTING SYSTEMS

The existing systems used to forecast weather require a lot of data about the current state of the weather and based on understanding of the atmospheric weather processes to predict how the weather evolves in the future. Some very sophisticated and complex methods which use satellite for prediction of weather exist. Other than such systems other systems which make use of machine learning techniques such as Linear regression and statistical methods also exist. The systems which have been proposed are having certain limitations. In some of the models, data of Weather is taken from the surroundings of the target place., More the surrounding data, less the efficiency as target weather does not always depend on its surrounding places. Over prediction is seen in many models. The predicted readings are 10%(on average) higher than the original values. Cost is high when we go for Satellite image processing which indeed gives high accuracy.

3. PROPOSED MODELS AND METHODOLOGY

To predict the weather conditions for the near future at a specific location with high accuracy and efficiency. The predicted weather conditions should match the actual weather conditions. To reduce the reliability on expensive instruments and methods presently used to predict the weather. And help daily commuters, farmers and other people by informing the weather situations before hand so that they can plan their schedule accordingly
To overcome the limitations and to reduce complexity, we have proposed two machine learning models :

1. Multi-target regression model
2. Recurrent Neural Network Model

3.1 MULTI-TARGET REGRESSION MODEL :

Multi-target regression is helpful in predicting simultaneously multiple continuous target variables based on the same set of input variables. Multi-target Regression Model is also known as Multivariate Regression or Multi-Output Regression.

The training data for this model is labeled data.
If the Training data is {(x1, y1),(x2, y2), . . . ,(xm, ym)}, $y_i \in R^q$ then the predict vector $y = \{y1, y2, \ldots, yq\}$ for a given input vector x.

For a feature vector x predict accurately a vector of responses by using a function h(x):
$X \rightarrow^{h(x)} y = (y1, y2, \ldots, yq)$

In the regression model, the input vector x could be parameters such as max-temperature,min-temperature, of the day, humidity percentage, overall weather such as sunny or cloudy.

The input variables are mapped to the multiple outputs that we want to predict using multi target regression. The outputs could be precipitation rate,wind speed and pressure.

Since this is a regression problem, training set includes all the labeled values of both inputs as well as outputs. When a new set of data for testing is presented to the model to predict the output, it will give out multiple values as outputs.

	A1	A2	Y1	Y2	.....	Y3
X1	6.0	6.5	7	5.3	.....	5.2
X2	9.0	2.8	14	4..2	.....	4.5
Xn	8.0	3.5	20	0.8	.....	3.3
X	3.0	7.5	?	?		?

3.2 RECURRENT NEURAL NETWORK (RNN) MODEL :

Recurrent neural network is a deep learning model which has an internal memory which can be used to store the information of the previously calculated output. In an RNN model, data goes through a loop and when the output is calculated, it considers the current input and also the input that was used in the previous iterations to improve the accuracy of the model.

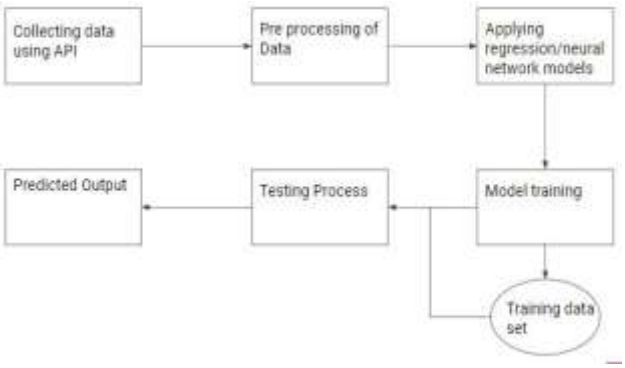
When there is a long sequence, since RNN has a short term it is very hard for the model to carry the information from the earlier time to the next.

RNN may leave out important information from the earlier parts.

Long Short–Term Memory (LSTM) networks are an extension of recurrent neural network, which makes it easier to recall past data in memory. It also helps in extending the memory of the RNN model. Vanishing gradient problem in the RNN is resolved by using the LSTM algorithm.

Here the proposed weather forecasting model using the RNN with Long Short Term Memory algorithm intends to gather data that is weather parameters, like temperature, humidity, pressure, dew point, wind speed, precipitation and visibility.

These are considered as neurons of input to recurrent neural networks. Weather forecasting is done by collecting information related today weather in regards to the previous and the present condition of the weather and utilizing this information to train LSTM model.



4. SYSTEM ARCJITECTURE:

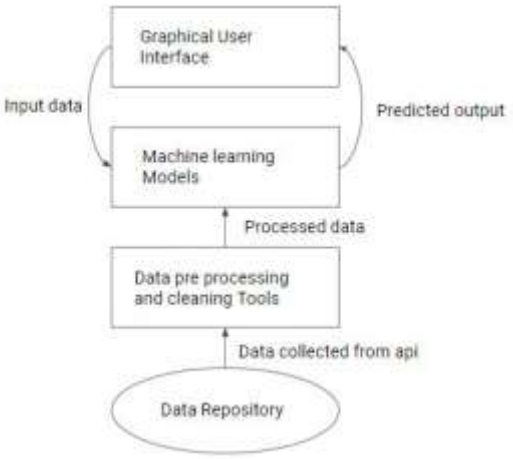
The complete system consists of a Graphical User Interface. The GUI takes the current day's weather input. These inputs are given to the trained models in order to predict the weather conditions of the near future.

The architecture of the system consists of a data repository consisting of the historical weather data of Bangalore city. This data is used for training the models.

Then it consists of data preprocessing tools in order to select the required data and to clean the data to get efficient results.

On top of that are the machine learning models which take in the preprocessed data and training the model to predict the future weather conditions.

And the topmost layer is the Graphical User Interface which helps user to provide inputs for the day and get accurate predictions.



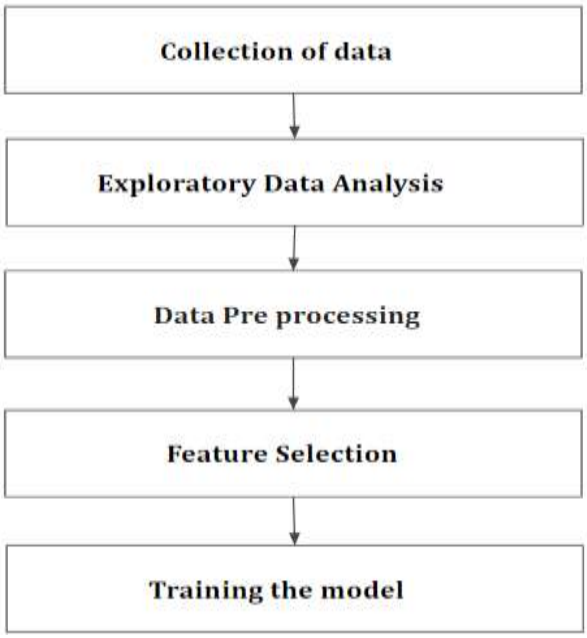
4.1 DATA FLOW:

Since data is the key component in machine learning models, it is very important to clean the raw data and extract only the required data and feed the data to the model in a specific format. The raw data obtained from the API has to be put through a series of steps for processing before it is used for training.

The first step is collection of historical data from a trusted source which is of high quality and precision. EDA is the process of conducting initial analysis of data in order to search for any existing helpful patterns, to test assumptions and hypothesis on the data and to find if any anomalies exist in the data. This is usually done with some visual and graphical representation. Some data mining tools are used to transform the raw

data into the desired format. Techniques such as data cleaning, data transformation, data reduction are used here to get the desired data. Not all the existing features are necessary for training the model. Only those features which contribute most to the accuracy of the model are selected here.

After selecting the required features the data is divided into a training set and testing set. The training set is used to train the machine learning model.



5. CONCLUSIONS

With the help of more advanced techniques of machine learning, we can only try to forecast the weather conditions. But we are not sure of the results matching exactly the same as the actual values, this is because weather also depend on the increase in the number of buildings and concrete structure, changes in the vegetation, an increase in the number of vehicles and pollution level. But these factors are a machine learning problem in their own domain. Integrating all these factors would be a challenging task. Utilization of machine learning models in prediction of weather conditions in short periods of time, which can run on less resource-intensive machines.Accurate prediction of weather related information in and around the city of Bengaluru.

Prediction of variables such as rainfall possibilities, temperature, humidity etc.

Evaluation of the proposed techniques and comparison of several machine learning models in the prediction of future weather conditions.

REFERENCES

[1] A H M Jakaria, Md Mosharaf Hossain, and Mohammad Ashiqur Rahman. 2018. Smart Weather Forecasting Using Machine Learning: A Case Study inTennessee. In Proceedings of ACM Mid-Southeast conference (Mid-Southeast'18). ACM, New York,NY, USA.

[2] Rohit Kumar Yadav and Ravi Khatri. 2016. A Weather Forecasting Model using the Data Mining Technique. International Journal of Computer Applications 139,14 (2016)

[3] Man Galih Salman, Bayu Kanigoro and Yaya Heryadi. 2015. Weather Forecasting using Deep Learning Techniques. ICACSIS 2015

[4] T.R.V.Anandharajan,G.AbhishekHariharan, K.K.Vignajeth, R. Jjijendiran Kushmita. 2016. Weather monitoring using Artificial Intelligence. 2016 International Conference on Computational Intelligence and Networks

[5] Tanzila Saba, Amjad Rehman, Jarallah S. A lGhamd,2017, Weather forecasting based on hybrid neural model . Springer Link

[6] Abrahamsen, Erik & Brastein, Ole & Lie, Bernt. (2018). Machine Learning in Python for Weather Forecast based on Freely Available Weather Data.

[7] Fente, Dires & Singh, Dheeraj. (2018). Weather Forecasting Using Artificial Neural Network.

[8] Chen, I-Ching & Hu, Shueh-Cheng. (2019). Realizing Specific Weather Forecast through Machine Learning Enabled Prediction Model. The 2019 3rd High Performance Computing and Cluster Technologies Conference

[9] Salman, A.G. & Heryadi, Y. & Abdurahman, E. & Suparta, Wayan. (2018). Weather forecasting using merged long short-term memory model. Bulletin of Electrical Engineering and Informatics.

[10] Lee, Jaedong & Hong, Sora & Lee, Jee-Hyong. (2014). An efficient prediction for heavy rain from big weather data using a genetic algorithm. Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication.