WEATHER PREDICTION AND CLIMATE ANALYSIS USING MACHINE LEARNING

Christina Mary Jolly

Post graduate student, *C.S.E Department, I.E.S College of engineering, Kerala, India*

ABSTRACT

Weather is an important aspect of a person's life as it can help us to know when it'll rain and when it'll be sunny. Weather forecasting is the attempt by meteorologists to predict the weather conditions at some future time and the weather conditions that may be expected. The climatic condition parameters are based on the temperature, pressure, humidity, dewpoint, rainfall, precipitation, wind speed and size of dataset. Here, the parameters temperature, pressure, humidity, dewpoint, precipitation, rainfall is only considered for experimental analysis.precipitation, wind speed and size of dataset. Here, the parameters temperature, pressure, humidity, dewpoint, precipitation, rainfall is only considered for experimental analysis.Weather forecasting is simply the prediction of future weather based on different parameters of the past like temperature, humidity, dew, wind speed and direction, precipitation, Haze and contents of air, Solar and terrestrial radiation etc. Weather forecast is an important factor affecting people's lives. Once the data is taken, it is trained. The heart of this project is the Linear Regression algorithm which is used to predict the weather using these data. The more parameters considered, the higher the accuracy. This project can help many people finding the weather of tomorrow.Prediction requires accurate classification of data .In order to predict the uncertain things, we need to analyse various factors which involved either directly or indirectly. Weather is one of the most influential environmental constraints in every phase of our lives on the earth. So as to make everyday tasks we are very much rely on weather and need to know weather condition on before hands. This could be achieved by predicting the weather condition such as humidity, rainfall, temperature, thunder, fog, etc. This helps us in protecting ourselves from abnormal conditions and avoids unnecessary delays. The main objective of this paper is to design an effective weather prediction model by the use of multivariate regression or multiple linear regressions and support vector machine (SVM). As of now, there are various debates going on around the world either scientifically or non-scientifically regarding the change of Earth's climate in fore coming decades/centuries and what impact it will cause on all the living creatures. Scientific models which predict future climates offer the best plan or aspiration for providing the information which will allow the world's policy maker to take preventive measures and make better decisions for the future of the Earth and for the future lives. This paper explores about weather forecast in effective way.

Keyword : - *Linear Regression , Naive b aye's classifier, Coefficient correlation, and Data preprocessing etc….*

1. INTRODUCTION

Weather forecasting is basically the prediction of the future weather and for the specified geographical location. Weather conditions are changing very rapidly around the world and it affects all the major areas. Weather forecasts become very essential in today's world. Today we are heavily depend on weather forecasts whether it is from industrial to agriculture, from travelling to daily commuting, anything where weather plays a role. For the easy and seamless mobility it becomes very important that we predict the weather correctly and ensure that it has no error.

The Scientists are still in working process of overcoming the limitations of computer models to improvise the accuracy rate of prediction through recent technologies of adding intelligence to machine. To add intelligence for system as human we have given a study platform called Artificial Neural networks, Machine learning, rule based techniques where there exist ample impetus to study the weather occurrence and prediction.

Here we implemented machine learning algorithm is Scikit's learn linear regression model. A large historical datasets of Austin was obtained and used to train this algorithm. The input was the weather conditions of Austin of many years and thus the output obtained is weather condition of upcoming days.

2. LITERATURE REVIEW

There are many research papers that have been published related to predicting the weather. A paper was published on 'The Weather Forecast Using Data Mining Research Based on Cloud Computing' This paper proposes a modern method to develop a service oriented architecture for the weather information systems which forecast weather using these data mining techniques. This can be carried out by using Artificial Neural Network and Decision tree Algorithms and meteorological data collected in Specific time. Algorithm has presented the best results to generate classification rules for the mean weather variables. The results showed that these data mining techniques can be enough for weather forecasting. Another paper was published on 'Analysis on The Weather Forecasting and Techniques' where they decided that artificial neural network and concept of fuzzy logic provides a best solution and prediction comparatively . They decided to take temperature, humidity, pressure, wind and various other attributes into consideration.

Another research paper titled 'Issues with weather prediction' discussed the major problems with weather prediction. Even the simplest weather prediction is not perfect. The one-day forecast typically falls within two degrees of the actual temperature. Although this accuracy isn't bad, as predictions are made for further in time. For example, in a place like New England where temperatures have a great variance the temperature prediction are more inaccurate than a place like the tropics. Another research paper titled 'Current weather prediction' used numerical methods to stimulate what is most likely going to happen based on known state of the atmosphere. For example, if a forecaster is looking at three different numerical models, and two model predict that a storm is going to hit a certain place, the forecaster would most likely predict that the storm is going to hit the area. These numerical models work well and are being tweaked all the time, but they still have errors because some of the equations used by the models aren't precise.

3. SYSTEM DESIGN

The application of science and technology that predicts the state of atmosphere at any given particular time period is known as Weather forecasting. There is a many different methods to weather forecast. Weather forecast notices are important because they can be used to prevent destruction of life and environment. The weather forecasting methods used in the ancient time usually implied pattern recognition i.e., they usually rely on observing patterns of events. For example, it is found that the following day has brought fair weather; if the preceding day sunset is particularly red. However, all of the predictions prove not to be reliable.

3.1 PROPOSED SYSTEM

Firstly, the data is trained. For training the data, we will take 15-20% of the data from the data set. For this prediction, we'll be using Linear regression algorithm and Naïve Bayesian classification algorithm. For the project, we'll be using python, NumPy, Jupiter Notebook, Spyder, Panda. The project is split into three separate Jupiter Notebooks: one to collect the weather data, inspect it, and clean it; a second to further refine the features and fit the data to a Linear Regression model and Naïve Bayesian model and a third to train and evaluate our output.

The project simply uses temperature, dew, pressure and humidity for training the data. Here these data are then trained using Linear Regression for the prediction.

The Naive Bayes algorithm is comprised of two words Naïve and Bayes, Which can be described as: Naive: It is called Naïve because it assumes that the occurrence of a certain feature is independent of the occurrence of other features. Such as if the fruit is identified on the bases of color, shape, and taste, then red, spherical, and sweet fruit is recognized as an apple. Hence each feature individually contributes to identify that it is an apple without depending on each other. Bayes: It is called Bayes because it depends on the principle of Bayes' theorem

Linear regression is one of the easiest and most popular Machine Learning algorithms. It is a statistical method that is used for predictive analysis. Linear regression makes predictions for continuous/real or numeric variables such as sales, salary, age, product price, etc

4. CONCLUSIONS

The weather prediction done using linear regression algorithm and Naïve Bayes algorithm are very essential for improving the future performance for the people. For predicting the weather, the linear regression algorithm and Naïve Bayes algorithm was applied to the datasets of the weather. We made a model to predict the weather using some selected input variables collected from Kaggle. The problem with current weather scenario is that we are not able to prepare our self and not able to do some important works. So, for knowing the weather scenario at high accuracy considering every factor that affects in the weather scenario, this model is

Therefore in this we provided how the machine learning techniques can be trained and used for the weather forecasting. In this Machine learning models are much accurate than human prediction and physical models prepared by human. For weather forecasting we used dataset for the Austin KATT station from 2013-02-21 to 2017-07-31. Accuracy obtained here was measured on the basis of coefficient correlation. We also utilize the historical data to predict the weather conditions which is much faster than the traditional models. The new pattern is combining deterministic and machine learning or statistical components, can provide fast and accurate calculations of these processes as well and help in predicting value of independent variable accurately. In future work, going to do research and make a model on how the neighboring weather can affect the weather of our area.

5 REFERENCES

[1]. Reference 1 Cohen, J., Cohen P., West, S.G., & Aiken, L.S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences. (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum Associates

[2]. Reference 2 Janani.B,Priyanka Sebastian. (2014). Analysis on the weather forecasting and technique. (3rd ed). IJARCET.

[3]. Reference 3 Samenow & Fritz. (2015). Issues with weather prediction.

[4]. Reference 4 Jabani B and Priyanka Sebastian. (2014). Analysis on The Weather Forecasting and Techniques.