Weather Forecasting Using Data Mining

Ankit Jaiswal, Prof. L.Vetrivendan Sir
School of Computer Science & Engineering Galgotias University, Greater Noida, UP

Abstract- **Climate forecasts are an important use of climate and have been one of the problems of science and technology worldwide for the past century. In this paper, we investigate the use of data mining techniques to predict high temperatures, precipitation, evaporation and wind. This was done using the random forest regressors using the decision tree algorithm and the mean of the output of the different different decision tree . A meteorological data model was developed and this was used to train classifier algorithms and the data is being drawn or taken from the country Sweden and the data is from 2000 to half of the 2019 . The performance of these algorithms was compared using standard performance metrics, and the algorithm yielded the best results used to extract the rules for the classification of such climate variables. The Neural Network prediction model has also been developed for the system for forecasting the weather and results in comparison to the actual weather data for the predicted times. The results indicate that if sufficient data are provided, Mining data strategies can be used for climate forecasts and climate studies.**

Index Terms- **Weather Forecasting, Data Mining, Neural Integration Networks, Decision Trees.**

I. INTRODUCTION

Climatic forecasts have been one of the most challenging scientific and technological challenges in the world in the last century. This is due to two factors: first, many human activities are used and secondly, due to emerging technological opportunities that are directly related to this field of concrete research, such as the emergence of integration and optimization of measurement systems.

Accurate prediction is one of the biggest challenges facing a global meteorologist. From ancient times, weather forecasts have been a fascinating and interesting background.

Scientists have tried to predict the climatic factors using a number of methods, some of which are more accurate than others.

Weather forecasts predict how the current climate will change. The current climate is achieved by ground observations, ship and aircraft observations, radiosondes, Doppler radar and satellites. This information is sent to environmental institutions where data is collected, analyzed, and organized into various maps, maps and graphs. High speed computers transfer thousands of views to more and more maps. Computers draw lines on maps with the help of meteorologists, who correct them for any errors. The final map is called an analysis. Computers not only map out maps but predict what maps will look like in the future. Computer

weather forecasts are known as weather forecasts in numbers.

Predicting the weather in numerical ways, meteorologists have developed atmospheric models that measure the atmosphere using statistics to explain how atmospheric temperature, pressure, and humidity will change over time. These figures are programmed into a computer and information on current weather conditions is captured on a computer. The computer solves those calculations to find out how the various wind changes will change in the next few minutes.

II. LITERATURE SURVEY

Climate change is an important and lasting change in the distribution of climate patterns statistics from decades to millions of years. It may be a change in the average weather conditions or the distribution of events around that average (e.g., more or less severe weather events). This term is sometimes used to refer directly to climate change caused by human activity, compared to climate change that may result as part of Earth's natural processes. Today's climate change is related to anthropogenic global warming. In scientific journals, however, global warming refers to rising temperatures, while climate change includes global warming and everything else that will be exacerbated by greenhouse gas prices. Evidence of climate change is taken from a variety of sources that can be used to rebuild past climate. Comprehensive global records of global warming were found from the mid-late 19th century. In earlier times, most of the evidence is indirect. Climate change is attributed to changes in proxies, climate indicators, such as vegetation, ice cores, dendrochronology, sea level changes, and glacial geology.

In 1988, the United Nations Environment Program and the World Meteorological Organization established the Intergovernmental Panel on Climate Change (IPCC) to examine the environmental, social, economic and scientific data available for climate change. The IPCC's Second Assessment Report, published in 1995, concluded that global air temperature increased by between 0.3 and 0.6 Celsius degrees (between 0.5 and 1.1 Fahrenheit degrees)
100 years agone. Their report says that this warming can continue and therefore the average world temperature can rise between one.0 and 3.5 Celsius degrees (between one.8 and 6.3 Gabriel Daniel Fahrenheit degrees) within the year 2100. within the event of this warming, ocean levels might increase between fifteen cm and ninety five cm (6 in and thirty seven in) by 2100, with the likelihood of a fifty cm (20 in) increase. Such a rise in water level will have a prejudicial result on the marine atmosphere.

Other changes {that can which will that may} occur as a results of this warming will embody changes in wind and world precipitation patterns. several climate scientists believe human activities ar answerable for warming. they're a significant explanation for warming by burning saturated fossil fuels, that increase the number of carbonic acid gas (CO2) within the atmosphere. carbonic acid gas levels currently gift at around 360 elements per million (ppm), have exaggerated by twenty eight % within the last century. the results, or impact, of temperature change will be physical, environmental, social or economic. it's foreseen that future climate changes can embody additional warming (that is, higher world warming), rising ocean levels, and a potential increase within the frequency of bound extreme weather events.

III. METHODS

(1) Data Collection

the information used for this work was collected from meteorologic department of the country Kingdom of Sweden. The case details cowl a amount, i.e. January 2000 to Dec 2019. the subsequent procedures were adopted for this analysis grade: information Purification, information Designation, information Transformation and Mining information.

(2) Data Cleaning

At this stage, the default info a model was developed that takes care of missing information, retrieve duplicate information, and take away unhealthy information. Finally, the refined information was reworked into an appropriate format for data processing.

(3) Data Selection

At this stage, the knowledge required for analysis was created a call and was out there on the dataset. The meteorologic info had 10 (5-6) signs, of that the kind and outline ar whereas within the applied math analysis of the numbers.

Due to the character of cloud information info wherever all of them values ar identical and a better share of losses the number of daylight information each within the analysis.

(4) Data Transformation

This is additionally referred to as information integration.

it's the introduce that the chosen information is reborn to a variable forms relevant to data processing. {the information the info the information} file was there saved in Commas Separated worth (CSV) file and databases were customised to reduce the result to live data.

(5) Data Mining Stage

The mining section is split into 3 phases. At every stage all algorithms ar accustomed analyze {the information|the info|the information} for the climate data. The check methodology obtained by this study was to divide by train by share of information, cross it and so by the remaining share.

IV. RESULT

In this section, we tend to show the experimental results of the planned model during which the no. of the error and therefore the rules generated and therefore the variety of tree is generated and every one the information is been shown within the tables given below.

The random forest regressor's graph is be shown below:

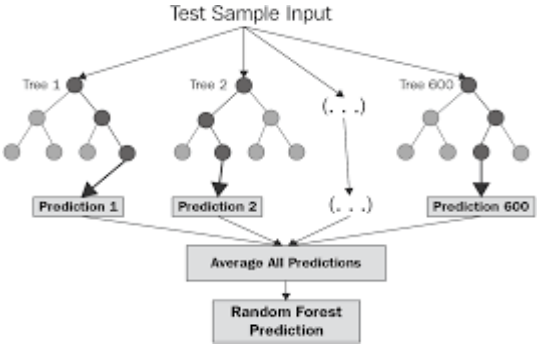

Fig. (1) – Test Sample Input

TABLE -1 ERROR DETECTION

Run No	No of Trees Generated	Error
1	21	58.3%
2	19	50.0%
3	21	41.7%
4	18	41.7%
5	16	753.0%
6	17	33.3%
7	15	58.3%
8	21	41.7%
9	18	58.3%
10	17	50.0%
Mean	18.3	50.8%
SE	0.7	3.8%

Table – 2 ERROR DETECTION

Run No	No of Rules Generated	Error
1	13	58.3%
2	16	50.0%
3	14	41.7%
4	16	50.0%
5	16	33.3%
6	13	58.3%
7	16	25.0%
8	17	33.3%
9	15	33.3%
10	20	41.7%
Mean	15.6	42.5%
SE	0.7	3.6%

A summary of the generation of See5 rules generation in experimental data using the 10-factor verification is presented in Table 2 of the rules from Run No. 7 that had the least errors identified.

After all the output from the form the decision trees we get finally we are now able to predict the future value of the temperature as well as the previous temperature value.

V. CONCLUSION

In this work the C5 tree choice algorithmic rule and random forest regressor is employed to get trees and selections to classify climate parameters like warmth, vasoconstrictor, rainfall, evaporation and wind speed counting on the month and year. the information used were of the urban center metropolis obtained from the meteorologic web site. The results show however these parameters have influenced the weather discovered throughout these months throughout the study amount. With the supply of comfortable information the discovered trend over time will be studied at the side of important deviations indicating changes within the discovered climate patterns.

Artificial Neural networks will find relationships between input variables and turn out results supported internal visual information patterns while not the requirement for programming or develop advanced calculations to simulate these relationships. thus given enough information ANN's will find the link between the climate parameter and use it to predict future weather. each TLFN neural networks and latent network structures ar accustomed extend the ANN models for predicting future values of wind speed, evaporation, radiation, vasoconstrictor, most temperature and precipitation provided by the Moon and therefore the year.

REFERENCES

□□□□https://opendata-download-metobs.smhi.se/api/version/1.0/parameter/2/station/71420/p eriod/corrected-archive/data.csv

□□□□□https://opendata-download-metobs.smhi.se/api/version/1.0/parameter/2/station/71420/p eriod/latest-months/data.csv

□□□□□http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6663175&queryText%3DWeather+Forecasting+Using+ Data+Mining

□□□□□http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=4136992&queryText%3DWeather+Forecasting+Using+ Data+Mining

□□□□□Bregman, J.I., Mackenthun K.M., 2006, Environmental Impact Statements, Chelsea: MI Lewis publication.

[6] Due R. A., 2007, A applied math Approach to Neural Networks for Pattern Recognition, eighth edition. New York: John Wiley and Sons publication.

[7] Elia G. P., 2009, "A call Tree for Weather

Prediction", Universitatea Petrol-Gaze din Ploiesti, Bd. Bucharesti thirty-nine, Ploiesti, Catedra Delaware Delaware, Vol. LXI, No. 1.

□□□□□Han, J., Micheline K., 2007, information Mining: ideas and Techniques, San Fransisco, CA: Morgan Kaufmann publishers.

□□□□□Martin T. H., Howard B. D, Mark B., 2002, Neural Network style, Shanghai: Thomson Asia PTE LTD and China Machine Press.

□□□□□ https://elitedatascience.com/python-machine-learning- tutorial-scikit-learn

□□□□□□https://nbviewer.jupyter.org/github/srnghn/ml_example_notebooks/blob/master/Predicting%20Yacht%20Resistance%20with%20Decision%20Trees%20%26%20Random%20Forests.ipynb?source=post_page

□□□□□□https://machinelearningmastery.com/make-predictions- scikit-learn/

[13] Wikipedia, 2010, "Effects of worldwide Warming" From Wikipedia - the free reference work, retrieved from http://en.wikipedia.org/wiki/Effects_of_Global_Warming in March 2010

□□□□□□ Wikipedia, 2011, "Climate change" From Wikipedia - the free reference work, retrieved from http://en.wikipedia.org/wiki/Climate_change in August 2011.