Weather Forecasting and Climate Changing Using Data Mining Application

Ankita Joshi[1], Bhagyashri Kamble[2], Vaibhavi Joshi[3], Komal Kajale[4], Nutan Dhange[5]

Student, Pursuing B.E., Atharva College of Engineering, Mumbai, India [1,2,3,4]

Professor, IT Department, Atharva College of Engineering, Mumbai, India [5]

Abstract: Weather forecasting is an important application in meteorology and has been one of the most scientifically and technologically challenging problems around the world. In this paper, we investigate the use of data mining techniques in forecasting maximum temperature, rainfall, evaporation and wind speed. Weather prediction approaches are challenged by complex weather phenomena with limited observations and past data. Weather phenomena have many parameters that are impossible to enumerate and measure. Increasing development on communication systems enabled weather forecast expert systems to integrate and share resources and thus hybrid system has emerged. Even though these improvements on weather forecast, these expert systems can't be fully reliable since weather forecast is main problem.

Keywords: Weather Prediction, Data Mining, Weather Conditions, Decision Trees

I. INTRODUCTION

Weather is one of the most effective environmental constraints in every phase of our lives. We are subject to adjusting ourselves with respect to weather condition from our dressing habits to strategic organizational planning activities, since the adverse weather conditions may cause a considerable damage on our lives and properties. We need to be on alert to these adverse weather conditions by taking some precautions and using prediction mechanisms for early warning of hazardous weather phenomena.

In this survey we try to give readers an overview about weather prediction phenomena, expert systems approaches, main domain specific problems, and solution methodologies.

A. Weather Prediction Problem Characteristics:

Weather prediction is a complicated procedure that includes multiple specialized fields of expertise. "There are only two methods to predict weather: the empirical approach and the dynamical approach" (Lorenz 19) [1]. Lorenz thus separated weather forecasting methodologies into two main branches in terms of numerical modelling and scientific processing (AI) of meteorological data.

B. Forecast Types:

Weather forecast expert systems can be categorized as terminal aerodrome forecasts (TAFs), public forecasts and marine forecasts in terms of field of interest and scale. In addition to these, many specialized forecasts are produced for specific purposes (e.g. weather pollution forecasting, agricultural predictions) TAFs are required to be most precise both in terms of measurable weather conditions and in terms of timing.

C. Importance of TAF

Airports organize air traffic mainly based on local weather condition that has agile parameters that may change in a short time. These parameters such as horizontal visibility and vertical visibility (ceiling), fogginess, precipitation level etc. can be very serious threat for flight safety and cost. When ceiling and visibility at a busy airport are low, in order to maximize safety, the rate of planes landing is reduced. When ceiling and visibility at a destination airport are forecast to be low at a flight's scheduled arrival time, its departure may be delayed in order to minimize traffic congestion and related costs. For Airport weather forecast timeliness and correctness are the main issues. An examination of the causes and effects of flight delays at the three main airports serving New York City concluded that a correctly forecast timing of a ceiling and visibility event (i.e., a significant change) could be expected to result in a savings of approximately $480,000 per event at La Guardia Airport.

II. RELATED WORK

Since the early 1950's (Tanner, 1952), it has been recognized that weather conditions affect driver behaviour and the manner in which a transportation system needs to be operated. By modifying speeds, headways as well as other parameters, drivers' reactions impact the overall system performance. This section presents a detailed literature review on the classification of inclement weather conditions and their translation into measurable objective parameters. The impact of such conditions on speed-flow-density relationships is first introduced. Such impact is associated with a change in capacity, delay, volume and speed, and reflects drivers' behaviour on a given road section. Once the change in these parameters is better understood, the control aspect of the study is analysed; research studies linking weather effects to signal timing, unsignalized intersections and variable message signs (VMS) are reviewed.

A. Weather Conditions:

The impact of "weather conditions" on transportation systems is a general term that may pose some confusion. Researchers have used different classification schemes for

weather conditions, because these conditions differ considerably in type and in magnitude. Some weather conditions are extreme in nature and thus may trigger a different response by the drivers.

Such extreme conditions are outside the immediate focus of the present study. Other inclement weather conditions (light and heavy rain, light and heavy snow etc.) offer a less compressed time frame to the decision makers, and allow drivers to retain an acceptable amount of control on their vehicles;

this control may be less than under "normal everyday" situation due to physical factors such as visibility, physical discomfort (cold or hot temperatures) and reduced pavement friction with the tires when there is precipitation or icy conditions prevail.

Most existing studies do not describe all "weather conditions" in the form of measurable objective parameters, making it difficult to explain or quantify the effect of such conditions on the transportation systems and their users.

Martin et al. (2000) suggested that before analysing the impact of such conditions, four dimensions need to be considered:

1. Severity of the condition
2. Duration
3. Geographic area of influence
4. Traffic flow or the demand served by the network

TABLE I
Rain Effects on Speed

Table 1. Rain Effects on Speed			
Researcher	**Ibrahim and Hall**	**Kyte et al.**	**Smith et al.**
Location	Toronto, Ontario	Idaho	Hampton Roads, Virginia
Year	1994	2001	2004
Speed Reduction in Light Rain	1.9-12.9 km/hr (1.2-8 mph)	9.5 km/hr (5.9 mph)	3-5%
Speed Reduction in Heavy Rain	4.8-16.1 km/h (3-10 mph)	9.5 km/hr (5.9 mph)	3-5%

TABLE II
Snow Effects on Volume

Table 2. Snow Effects on Volume			
	Freeway	**Arterial**	
Researcher	Hanbali and Kuemmel	Knapp	Maki
Location	Illinois, Minnesota, New York, Wisconsin	Iowa	Minneapolis, Minnesota
Year	1992	1995-1998	1999
Volume Reduction in Light Snow	7-29%	-	-

TABLE III
Weather Impact on Macroscopic Traffic Parameters

Table 3. Summary of Weather Impact on Macroscopic Traffic Parameters				
Factors\ Reduction	**Volume**	**Maximum Observed Flow**	**Capacity**	**Speed**
Rain	-	0-20%	4-47%	-
Snow	7-47%	5-10%	30%	13-40%
Wind	-	-	-	9%

III.   APPLICATION

The main idea is to choose the variable that provides more information to realize the appropriate partition in each branch in other to classify the training set. One advantage of Decision Tree classifiers is that rule can be inferred from the trees generated that are very descriptive, helping users to understand their data.See5 software can generate both decision trees and decision tree rules depending on selected options. The Trees and rules were generated using 10 fold cross validation and the results with the least error on the test data set were selected. Table 3 presents the summary of the runs and the decision tree obtained from Run Number 6 which had the least error.

Recurrent networks are the state of the art in nonlinear time series prediction, system identification, and temporal pattern classification and are of two types. Fully recurrent networks feedback the hidden layer to itself and Partially

recurrent networks start with a fully recurrent network and add a feed forward connection that bypasses the recurrence, effectively treating the recurrent part as a state memory. These recurrent networks can have an infinite memory depth and thus find relationships through time as well as through the instantaneous input space. Most real-world data contain information in its time structure.

Both the fully recurrent and partial recurrent network types were used.

IV.   PROPOSED SYSTEM

Weather forecasting entails predicting how the present state of the atmosphere will change. Present weather conditions are obtained by ground observations, observations from ships and aircraft, radio sondes, Doppler radar, and satellites. This information is sent to meteorological centres where the data are collected, analysed, and made into a variety of charts, maps, and graphs. Modern high-speed computers transfer the many thousands of observations onto surface and upper-air maps. Computers draw the lines on the maps with help from meteorologists, who correct for any errors. A final map is called an analysis. Computers not only draw the maps but predict how the maps will look sometime in the future. The forecasting of weather by computer is known as numerical weather prediction.

Weather forecasting has been one of the most scientifically and technologically challenging problems around the world in the last century. This is due mainly to two factors: first, it's used for many human activities and secondly, due to the opportunism created by the various technological advances that are directly related to this concrete research field, like the evolution of computation and the improvement in measurement systems. To make an accurate prediction is one of the major challenges facing meteorologist all over the world. Since ancient times, weather prediction has been one of the most interesting and fascinating domain. Scientists have tried to forecast meteorological characteristics using a number of methods, some of these methods being more accurate than others

To predict the weather by numerical means, meteorologists have developed atmospheric models that approximate the atmosphere by using mathematical equations to describe how atmospheric temperature, pressure, and moisture will change over time. The equations are programmed into a computer and data on the present atmospheric conditions are fed into the computer. The computer solves the equations to determine how the different atmospheric variables will change over the next few minutes.

CONCLUSION

In this project we used data mining algorithm we used Decision tree algorithm for classifying weather parameters such as maximum temperature, minimum temperature, rainfall, evaporation and wind speed in terms of the month and year. The data used from wounderground weather site between 2001 and 2013. The results show how these parameters have influenced the weather observed in these

months over the study period. Given enough data the observed trend over time could be studied and important deviations which show changes in climatic patterns identified.

The results obtained were evaluated with the test data set prepared along with the training data and were found to be acceptable considering the small size of the data available for training and testing. To have a better result a larger data set which will comprise of data collected over many decades will be needed. In future research works neuro-fuzzy models will be used for the weather prediction process. This work is important to climatic change studies because the variation in weather conditions in term of temperature, rainfall and wind speed can be studied using these data mining techniques.

REFERENCES

[1]   Application of Data Mining Techniques in WeatherPrediction and Climate Change Studies Published Online February 2012 in MECS (http://www.mecs-press.org/) DOI:10.5815/ijieeb.2012.01.07
[2]   Ahrens, C. D., 2007, "Meteorology" Microsoft®   Student 2008 [DVD], Redmond, WA: Microsoft Corporation, 2007.
[3]   Bregman,J.I.,Mackenthun K.M., 2006,Environmental Impact Statements, Chelsea: MI Lewis Publication.
[4]   Casas D. M, Gonzalez A.T, Rodrígue J. E. A., Pet J. V., 2009, "Using Data-Mining for Short Term Rainfall Forecasting", Notes in Computer Science, Volume 5518, 487-490
[5]   WEATHER PREDICTION EXPERT SYSTEM APPROACHES (Ceng-568 Literature Survey)
[6]   Incorporating Weather Impacts in Traffic   Estimation and Prediction Systems - ITS Report.html

.