Weather prediction using the data mining Techniques

Aishwarya Dhore[1], Anagha Byakude[2], Bhagyashri Sonar[3], Mansi Waste[4]

[1,2,3,4]DYP Pimpri College of Engineering, Pune, Maharashtr,India

---***---

Abstract - *Data mining techniques has been the subject of several research papers. Climate change may have a serious impact on the availability especially in countries of human resources. Knowledge discovery from temporal, spatial and spatiotemporal data is critical for climate change and climate effect on the environment. Climate change will significantly impact, public health. Historical climate statistics for future forecasting. The evolution of the tools and techniques available to gather data about weather, water, temperature, etc. In order to prevent these public health impacts, we can take actions to prevent climate change by analyzing historical data. Climate change refers to long-term shifts in weather conditions and patterns of extreme weather events. It may lead to changes in health threat to human beings, multiplying existing health problems. To go beyond empirical observations of the association between climate changes, infections and develop system that will predicts the climate effect and what the preventive measure should have to take by classification and clustering of data. To improve the prediction of spatial temporal process of climate change and the associated effects related to the various geographical scales.*

***Key Words*:   Weather forecasting, data mining, and prediction algorithms, Numerical Weather Prediction models**

1. INTRODUCTION

The importance of understanding and improving the quality of the prediction system of climate is important. Fresh climate and the environmental conditions are the stream flow of these and availability of human, animal and the natural hazards may cause adverse effects on the human beings. Stream flow prediction therefore provides crucial information [8] for the Weather prediction system. Global climate change challenge scientists and engineers to estimate and forecast [8] and the all magnitude that will discharge the accuracy. Faults are defined as adverse situation with significant impact on climate [2]. Due to the unbalanced volume of the data there is fault in accurate weather prediction. Data can be handle by merging of data because lack of data. Weather prediction system is the application of science and technology to predict the atmosphere [6] for the location or specific region. Weather forecasting is the process of identifying the climatic conditions using data mining techniques. Many of the system may depend on the weather prediction system. Weather prediction system has been helpful to prevent the damage of life and large extent of property. Weather forecast like temperature humidity and about rainfall which are important in agriculture and other commodity market. We take example of temperature prediction used to many companies to calculate or estimate demand for coming days. Weather forecast can be used to plan activities and events plan further ahead and survive [2].

Now days there are many technique are used for the weather forecasting and improving accuracy of forecasting. There are multiple types of predictions includes in weather forecasting.

2. Background

In area o weather prediction system there are many researchers have tried by using some data mining methodologies [4]. The K-means clustering is to divide the land and other areas to discover the interesting patterns. Dataset analyzed on the basis of clustering and classification. The probability of weather prediction is discussed. Weather is predicted based on the uncertainties of initial conditions and model formulation [4]. The accuracy of results is more important because on the weather forecasting because some peoples are rely on the site to manage their events. The reason depend on climate decision of events is taken. Work on the real time dataset normalized using min-max normalization [7]. Quantitative forecast like temperature, humidity and weathers conditions are important in agriculture area, to plan for the future and planning to trip in the city, as well as to traders within commodity markets. Weather and climate disasters are on the rise in the India no national capacity to provide skillful long range severe weather outlooks. Many variables that may affect weather, considering all the current statistics on the climate, physical equations to describe particle interactions on the smallest scale are aggregated to model an entire weather system. To assess the current state of the science and to identify what is required to develop long range severe weather prediction system for the future prediction and the preventive measure to do. Forecasting helps to take necessary measures, to prevent damage to life and property to a large extent. Quantitative forecast like temperature, humidity are important in agriculture area, as well as to traders within commodity markets. Different categories of forecasting methods are Naive approach, Judgmental methods, Quantitative and Qualitative method, Causal or econometric forecasting methods.

3. RELATED WORK

Yuko Tachibana and Mikihiko Ohnari presents Model of Hourly Water Consumption in Water Purification Plant through Categorical Approach because the need of prediction of hourly water consumption while supplying water to consumers. Water consumption is shown in waveform and is influenced of the week weather and temperature they resembles each other. Water comsuption on holidays and other days are not same. By analyzing the water consumption in metropolitan cities and by applying data mining techniques constructed precise prediction for year [1].

Andrew Kusiak and Shital Shah proposed simple and efficient alarm system for predicting faults incoming in water chemistry. This system is based on the modular, data driven and data mining based. In this system function are used such as, preprocessing, learning, prediction, alarm generation, and display. A Decision making system used for predicting the water chemistry faults with alarm system. After alarm boiler will get shutdown [2].

Charles a. Doswell et.al, examines the method for flood forecasting. This is developed by using the notion of ingredients. It is the result of high rainfall rates. This contains water vapor in the air and also depends on water precipitation.

Described details of ingredients and recognition techniques for heavy precipitation [3].

Zhongnan Zhang et.al, describes the association rule from the datagram, in the large scale data to find the important patterns to by the forecaster to predict the severe of the weather situations. Individual analysis is performed for the weather analysis. Zhongnan Zhang et.al proposed new algorithm called as DIAL for finding relationship special weather change and sever of climates. There are three stages in the proposed algorithm first change static dataset by recording weather condition with the new dataset by changing tendency of weather for every measurement. Second apply association rule on the newly generated dataset. Last use predefined predicts for transferring mediator rule in dynamic inter-dimension association rules [4].

J.P. Evans tries to improve by presenting regional climate model for characterizing of streamflow. Give information related to the complex land surface parameterization into regional climate model. Focused on the improving the land modeling by taking feedback to the climate. Present model for improving in the characteristics of the of prediction by Catchment Moisture Deficit– Identification of unit Hydrographs And Component flows from Rainfall, Evaporation and Streamflow data CMD-IHACRES) [5].

P. S. Mohod et.al states that rainfall and weather forecasting is most challenging problem in around the world in the agricultural field. Describes the algorithms related to data mining which are used for the determination. Neural network (NN), random forest, classification and regression tree (CRT), support vector machine (SVM) and k-nearest neighbor. By default, these algorithms are used for prediction. On the available dataset apply frequent mining for discovering frequent pattern. Frequent item will find from dataset on parameters like temperature, humidity, wind. These algorithms applied on the last five year rainfall dataset. Predict rainfall or not [6].

Priyanca Fargose et.al reviewed for a better approach for weather prediction and examines Artificial Neural Network, Ensemble Neural Network, Backpropagation Network, Radial Basis Function Network, General Regression Neural Network, Genetic Algorithm, Multilayer Perceptron, Fuzzy clustering, etc. Mentioned algorithm used for different types of weather prediction. Backpropagation algorithm is best for weather prediction. Neural network entitled as todays automatic itself generate data and predict the climate conditions by analyzing data. The number of parameters can   be taking into consideration while predicting weather of the day [7].

A. Makkeasorn et.al, given comparative study on the global climate change main between neural network and genetic algorithm. Study contains related to sea surface temperature and location based rainfall details via Next Generation Radar (NEXRAD). Study related to the artificial intelligence approaches, has two parts neural network and genetic algorithm. This work assessed with the predicting correct weather conditions by analyzing SST and metrological data [8].

G. Atsalakis and C. Minoudaki, presented the way to schedule the water pumping efforts and trying to minimize the cost required to it. Examines the demand of water consumer and predict the water distribution. Adaptive Neuro-Fuzzy Inferences System (ANFIS) technique used for itself prediction of water distribution by mining last stored details [9].

Ayham Omary et.al, studied precipitation and weather information by using Artificial Intelligence and data mining techniques. Tool is used to store and parse stored data. Parsed data feed to AI engine and data mining for further future weather change and precipitation information from historical data [10].
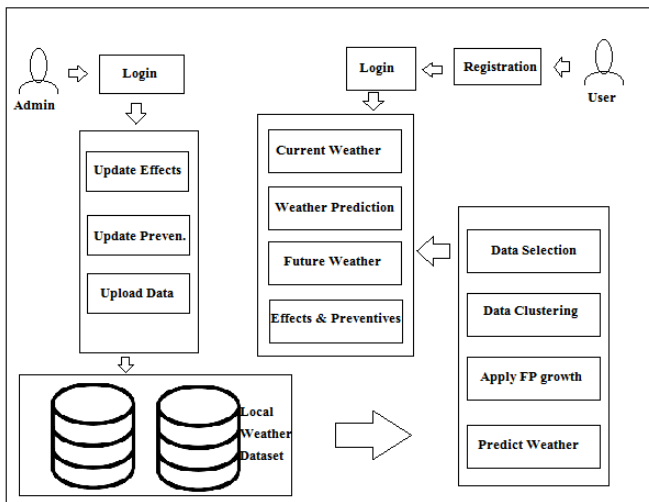
4. PROPOSED SYSTEM

Fig. Proposed Architecture

In the paper we propose the weather prediction system which predicts the current weather conditions and the future weather conditions. The data mining techniques are applied on the dataset to extract the useful information from the dataset. Such clustering of data, to find the frequent item set extraction and the genetic algorithm for the best fitness of the weather conditions, effects and preventive measures. Proposed system shows in the above figure. The role of the admin is to upload the data such as effect of the temperature and preventive measure and upload the dataset to system. On the other hand at a client side user have to register to the application. After the login user gets the current temperature prediction and

Algorithm: FP Growth FP-Growth reads dataset transaction at a time and maps it

1. Fixed order is used, so paths can overlap when transactions share items

   – In this case, counters are incremented

2. Relations are maintained between nodes containing the same item, creating singly linked list.

Algorithm1: FP Growth Algorithm

Future prediction with the effects of the temperature along with preventive measures. To predict the weather conditions in the proposed system we used the data mining algorithms. To partition the data and to find the temperature or weather condition Clustering kmean and FP Growth algorithms used .

After that final predicted temperature input to find the best suitable effects and the preventive measure using the genetic algorithm.

5. DATA COLLECTION

Data collection is the process gathering of the required data for the processing purpose. We have collected from the locale city center and crawled from the open weather site for the particular region. The data has been collected of the last few years. We have taken few parameters such as temperature, wind humidity and weather conditions.

6. RESULT
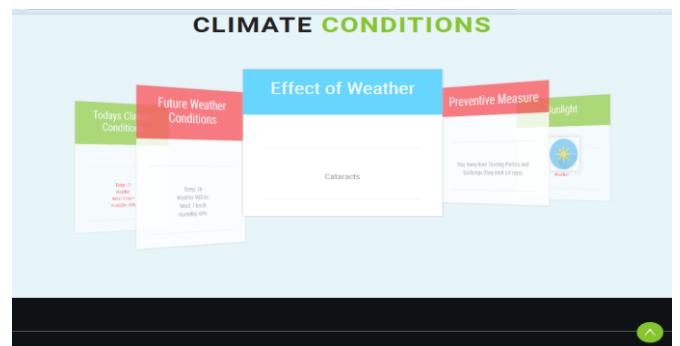

Figure. Predicted weather with temperature

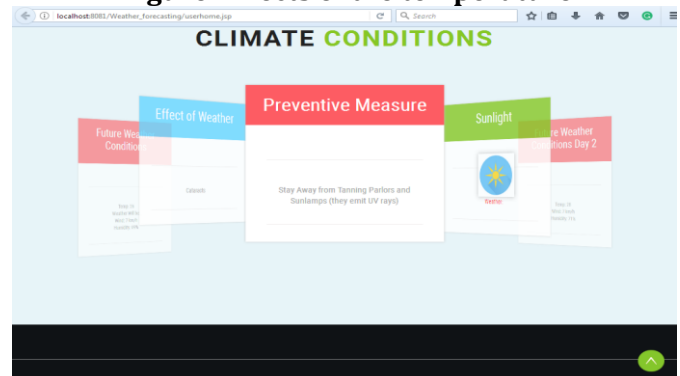**Figure. Effects of the temperature**

Figure. Preventives Measures

7. CONCLUSIONS

This system presents a first glance on a project to build a numerical weather prediction model in local weather. Such models are usually complex and will take significant time and resources to accomplish. Collect and have all required data and build a historical dataset of weather, precipitation and all possible related attributes. It accepts all complex parameters as input and generates the intelligent patterns while training and it uses the same patterns to generate the forecasts.

REFERENCES

[1] Prediction Model of Hourly Water Consumption in Water Purification Plant through Categorical Approach, Yuko Tachibana and Mikihiko Ohnari, 1999. IEEE SMC '99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics,

[2] Data-Mining-Based System for Prediction of Water Chemistry Faults, Andrew Kusiak, and Shital Shah, IEEE Transactions on industrial electronics, vol. 53, no. 2, april 2006.

[3] Flash Flood Forecasting: An Ingredients-Based Methodology, Doswell, C. A., et al, 1996: Weather and Forecasting, 11 560-581

[4] "Mining Dynamic Interdimension Association Rules for LocalscaleWeatherPrediction", Zhongnan Zhang, Weili Wu, Yaochun Huang, Proceedings of the 28th Annual International Computer Software and Applications Conference (COMPSAC'04), 2004

[5] Improving the characteristics of streamflow modeled by regional climate models. Evans, J.P., 2003. J. Hydrology 284, 211 – 227

[6] Amruta A. Taksande, P. S. Mohod Applications of Data Mining in Weather Forecasting Using Frequent Pattern Growth Algorithm

[7] Meera Narvekar, Priyanca Fargose Daily Weather Forecasting using Artificial Neural Network

[8] Short-term streamflow forecasting with global climate change implications – A comparative study between genetic programming and neural network models, A. Makkeasorn, Chang a, X. Zhou, Journal of Hydrology (2008) 352, 336– 354

[9] Daily irrigation water demand prediction using Adaptive Neuro-Fuzzy Inferences Systems (ANFIS), g. atsalakis and C. Minoudaki, proc. Of the 3rd iasme/wseas Int. Conf. on Energy, Environment, Ecosystems and Sustainable Development, Agios Nikolaos, Greece, July 24-26, 2007