iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VI **Month of publication:** June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35519

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

A Weather Forecasting Model using the Data Mining Technique

Tanvi Patil[1], Dr. Kamal Shah[2]
[1]*PG Student,* [2]*Professor and HOD, Department of IT*
Thakur College of Engineering and Technology, Kandivali (E), Mumbai,Maharashtra 400101

Abstract: ***The weather conditions are changing continuously and the entire world is suffers from the changing Clemet and their side effects. Therefore pattern on changing weather conditions are required to observe. With this aim the proposed work is intended to investigate about the weather condition pattern and their forecasting model. On the other hand data mining technique enables us to analyse the data and extract the valuable patterns from the data. Therefore in order to understand fluctuating patterns of the weather conditions the data mining based predictive model is reported in this work. The proposed data model analyse the historical weather data and identify the significant on the data. These identified patterns from the historical data enable us to approximate the upcoming weather conditions and their outcomes. To design and develop such an accurate data model a number of techniques are reviewed and most promising approaches are collected.***

Keywords: Machine learning, weather forecasting, prediction, Data Mining, Data pre-processing, Transformation.

I. INTRODUCTION

In recent years a number of new techniques and the new technologies are appeared. Among them the machine learning and data mining applications are producing their noteworthy contributions. In the various applications such as e-commerce recommendation, stock market prediction, spam filtering and others are developed with the help of data mining techniques. The data mining techniques are offered to analyse the historical data and prepare their experiences. This experience or the learning is used to identify the similar kind of data for classification task, or for making prediction and approximation. In this proposed work the machine learning based classification and prediction is studied in detail. In addition of that using the data mining technique a new prediction model is prepared using the hybrid technique of machine learning. The proposed technique is a weather forecasting model; the proposed weather forecasting technique is analysing the historical weather data and the concerned attributes. For evaluation of the attributes and identification of similar patterns the proposed technique includes both the supervised learning and unsupervised learning technique. Therefore the proposed technique is a hybrid technique of learning and prediction.

II. BACKGROUND

The backbone of Indian economy is Agriculture. Now a day's weather or rainfall is the stimulating problems around the world. Rainfall prediction is nothing but weather forecasting. Weather forecasting application is an art of science and technology use to the state of atmosphere for a location. The weather forecaster's work 24/7, 365 days of the year, using supercomputers it is easy to predict the weather for hours, days, weeks, seasons and even years ahead.

Weather forecasting is mainly concerned with the prediction of weather condition in the given future time. Weather forecasts provide critical information about future weather. There are various approaches available in weather forecasting, from relatively simple observation of the sky to highly complex computerized mathematical models. The prediction of weather condition is essential for various applications. Some of them are climate monitoring, drought detection, severe weather prediction, agriculture and production, planning in energy industry, aviation industry, communication, pollution dispersal, and so forth. In military operations, there is a considerable historical record of instances when weather conditions have altered the course of battles. Accurate prediction of weather conditions is a difficult task due to the dynamic nature of atmosphere. The weather condition at any instance may be represented by some variables. Out of those variables, one found that the most significant are being selected to be involved in the process of prediction. The selection of variables is dependent on the location for which the prediction is to be made. The variables and their range always vary from place to place. The weather condition of any day has some relationship with the weather condition existed in the same tenure of precious year and previous week. Rainfall is a form of precipitation. Its accurate forecasts can help to identify possible floods in future e and to plan for better water management.

In recent years a number of new techniques and the new technologies are appeared. Among them the machine learning and data mining applications are producing their noteworthy contributions. In the various applications such as e-commerce recommendation, stock market prediction, spam filtering and others are developed with the help of data mining techniques. The

data mining techniques are offered to analyse the historical data and prepare their experiences. This experience or the learning is used to identify the similar kind of data for classification task, or for making prediction and approximation.

III. MOTIVATION

A. Forecasting is a Vital Element

The World Meteorological Organisation says weather forecasting "is a vital element" needed "in order to meet the food, fodder, fibre and renewable agri-energy needs of rapidly growing populations".

Agriculture and farming are mainly dependent on seasons and weather. The temperature matters a lot in that case when it comes to the farming of different kinds of fruits, vegetables, and pulses. Previously we did not have a better understanding of weather forecasting and farmers were still doing their job based on predictions. Though sometimes they occur loss due to false predictions of weather. Now that the technology is developed and special weather forecasting mechanisms are available, the farmers can get all the updates are on a smartphone. Education towards that is, of course, an important thing but most of the farmer population at this stage knows the basics which make it easy for them to use the features.

Weather forecasting is a prediction on conditions of atmosphere depending on location and time. Every area will have their different predictions related to the condition of weather which makes pretty easy for the farmers to know how and what to do when. The relationship between weather and agriculture has, therefore, necessitated the need for accurate prediction of the weather; to enable farmers to make an informed decision that will not bring losses to them. Temperature, sunlight, and rainfall have major effects on the crops. For livestock, temperatures and adequate water and food are essential.

The forecast of the weather event helps for suitable planning of farming operations. It helps to decide whether to undertake or withhold the sowing operation. To irrigate the crop or not, when to apply fertilizer and whether to start complete harvesting or to withhold it are the major components for which forecasting is a must.

IV. PROBLEM STATEMENT

Weather forecasts are made by collecting as much data as possible about the current state of the atmosphere (particularly the temperature, humidity and wind) and using understanding of atmospheric processes (through meteorology) to determine how the atmosphere evolves in the future.

It is the task to observe meteorology data to determine the state of the atmosphere at agreed-upon, suitable time two years of and with sufficient accuracy.

Weather and climate over the earth are not constant with time; they change on different time series ranging from the geological to the diurnal through annual, seasonal and intra-seasonal timescales. Rainfall varies with latitude, elevation, topography, seasons, distance from the sea, and Coastal sea-surface temperature. Trends in rainfall extremes have enormous implications. Flood or draught events cause significant damage to agriculture, environment, infrastructure, social life and on GDP. Therefore, a reliable rainfall forecasting and assessing behavior at the station, regional and national levels is very important.

- **"**For forecasting the pattern of rainfall in the study area.
- To provide information that would be helpful for decision makers in formulating policies to mitigate the problems of rainwater resources management, soil erosion, flooding, and drought.
- To provide information for the early warning system in the study area. **"**

V. DATA MINING TECHNIQUES

Data mining is the process of exploring a large amount of data in order to find useful patterns for practical application. It is a powerful technology that uses machine learning, statistical and visualization techniques to discover and predict knowledge to the user. Prediction requires to develop a model able to discover relationships between dependent and independent variables. There are various types of data mining techniques such as: Association, Classification, Prediction, Text and Clustering. Data mining provides many tools by which big data can be analyzed automatically. Random forests, neural networks and Naïve Bayes are used in statistics, data analysis and machine learning.

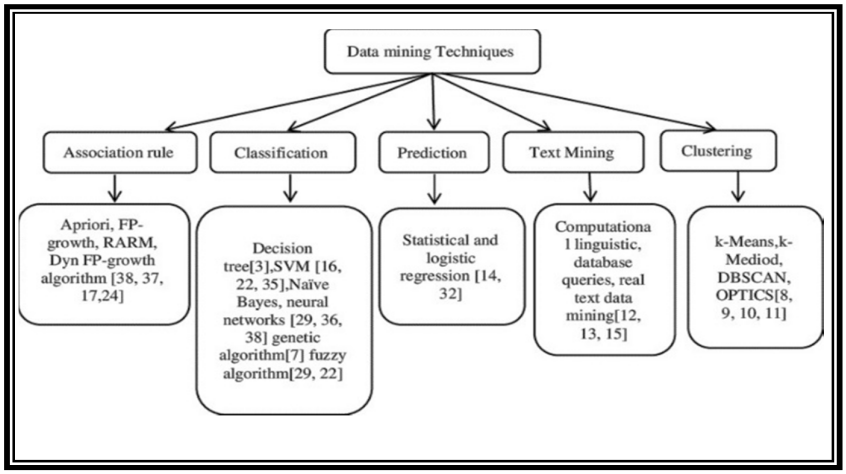


Fig 1: Classification of data mining techniques

Data mining is of two types, Descriptive data mining, and Predictive data mining. Descriptive mining describes concepts or task-relevant data sets in concise, summarized, informative, discriminative forms, and Predictive mining based on data and analysis, constructs models for the database, and predicts the trend and properties of unknown data. This paper analyses the effectiveness of predictive and classification algorithms for predicting rainfall and temperature.

VI. METHODOLOGY

A. System Model

A data flow diagram (DFD) is a graphical representation of the flow of data through an information system. A data flow diagram can also be used for the visualization of data processing (structured design). It is common practice for a designer to draw a context-level DFD first which shows the interaction between the system and outside entities. This context-level DFD is then exploded to show more detail of the system being modeled.

The four components of a data flow diagram (DFD) are:

1) External Entities/Terminators are outside of the system being modeled. Terminators represent where information comes from and where it goes. In designing a system, we have no idea about what these terminators do or how they do it.
2) Processes modify the inputs in the process of generating the outputs
3) Data Stores represent a place in the process where data comes to rest. A DFD does not say anything about the relative timing of the processes, so a data store might be a place to accumulate data over a year for the annual accounting process.
4) Data Flows shows how data moves between terminators, processes, and data stores (those that cross the system boundary are known as IO or Input Output Descriptions).

B. Methodology

The methodology used in the paper consists of certain steps that are usually used in data mining applications the steps are as follows:

1) Data Collection and Retrieval
2) Data Transformation
3) Data Pre-processing
4) Feature Extraction
5) Data mining

C. Proposed System

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. 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.

Phases of Weather Forecasting Process

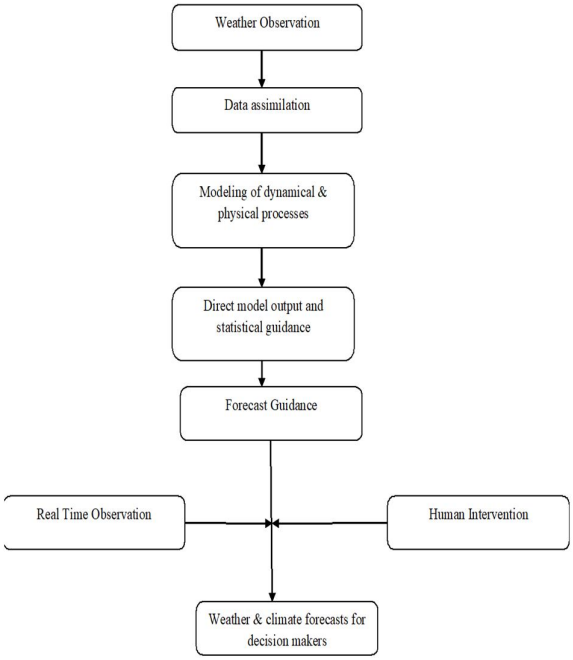


Fig 2: Phases of Weather Forecast

VII. RELATED WORK

The research methodology used in is Design Science Research Methodology (DRM).This method consists of 3 stages:
Data Understanding
Model Formulation
Evaluation



Fig 3: Research Workflow

A. Data Understanding

Start by collecting data, then get familiar with the data, to identify data quality problems, to discover first insights into the data, or to detect interesting subsets to form hypotheses about hidden information.

B. Data Preparation

Includes all activities required to construct the final data set (data that will be fed into the modelling tool) from the initial raw data. Tasks include table, case, and attribute selection as well as transformation and cleaning of data for modelling tools.

C. Modelling

Select and apply a variety of modelling techniques, and calibrate tool parameters to optimal values. Typically, there are several techniques for the same data mining problem type. Some techniques have specific requirements on the form of data. Therefore, stepping back to the data preparation phase is often needed.

D. Evaluation

Thoroughly evaluate the model, and review the steps executed to construct the model, to be certain it properly achieves the business objectives. Determine if there is some important business issue that has not been sufficiently considered. At the end of this phase, a decision on the use of the data mining results is reached.

E. Deployment

Organize and present the results of data mining. Deployment can be as simple as generating a report or as complex as implementing a repeatable data mining process.

VIII. WEATHER FORECAST METHODS

With the new advances in technology and data analytics techniques, people can be aware of extreme weather or any change in the atmosphere conditions, for example temperature, wind speed and direction, humidity, sunshine, cloudiness and precipitation. Deviations in these 18 conditions describe weather in terms of the state of the atmosphere at a specific location and time. According to Winston et al. (2018), weather forecasts generally follow three major steps:

- Observation and data collection of the atmosphere, the ocean, and land surface
- Assimilation, processing and analysis and extrapolation to predict the future state of the atmosphere.
- The totality of observations, analysis, model and computer system constitute a forecast system.

The different methods used in weather forecasts are: Synoptic weather forecasting, Numerical methods, and Statistical methods.

A. Synoptic Weather Prediction

Synoptic Weather Prediction is the traditional approach in weather prediction. Synoptic charts form the very basis of weather forecasts. Synoptic charts contain the analysis upon huge amounts of observational data of different weather elements collected from weather stations at a specific time. Meteorological centers prepare a series of synoptic charts every day to track weather changes, which forms the very basic of weather forecasts. From the careful study of weather charts over many years, certain empirical rules were formulated. These established rules have improved the forecasts estimating rate and direction of the movement of weather systems.

B. Numerical Weather Prediction (NWP)

Numerical Weather Prediction (NWP) is one of the most imperative operational tasks carried out by meteorological services around the world (Gorger et al., 2016). NWP applies computer algorithms to forecast weather. Supercomputers run complex computer programs to provide predictions on many atmospheric conditions. This method involves a set of partial differential equations and other formulations describing the dynamic and thermodynamic processes of the earth's atmosphere. It is the comprising of equations, numerical approximations, parametrizations, domain settings as well as initial and boundary conditions. 19

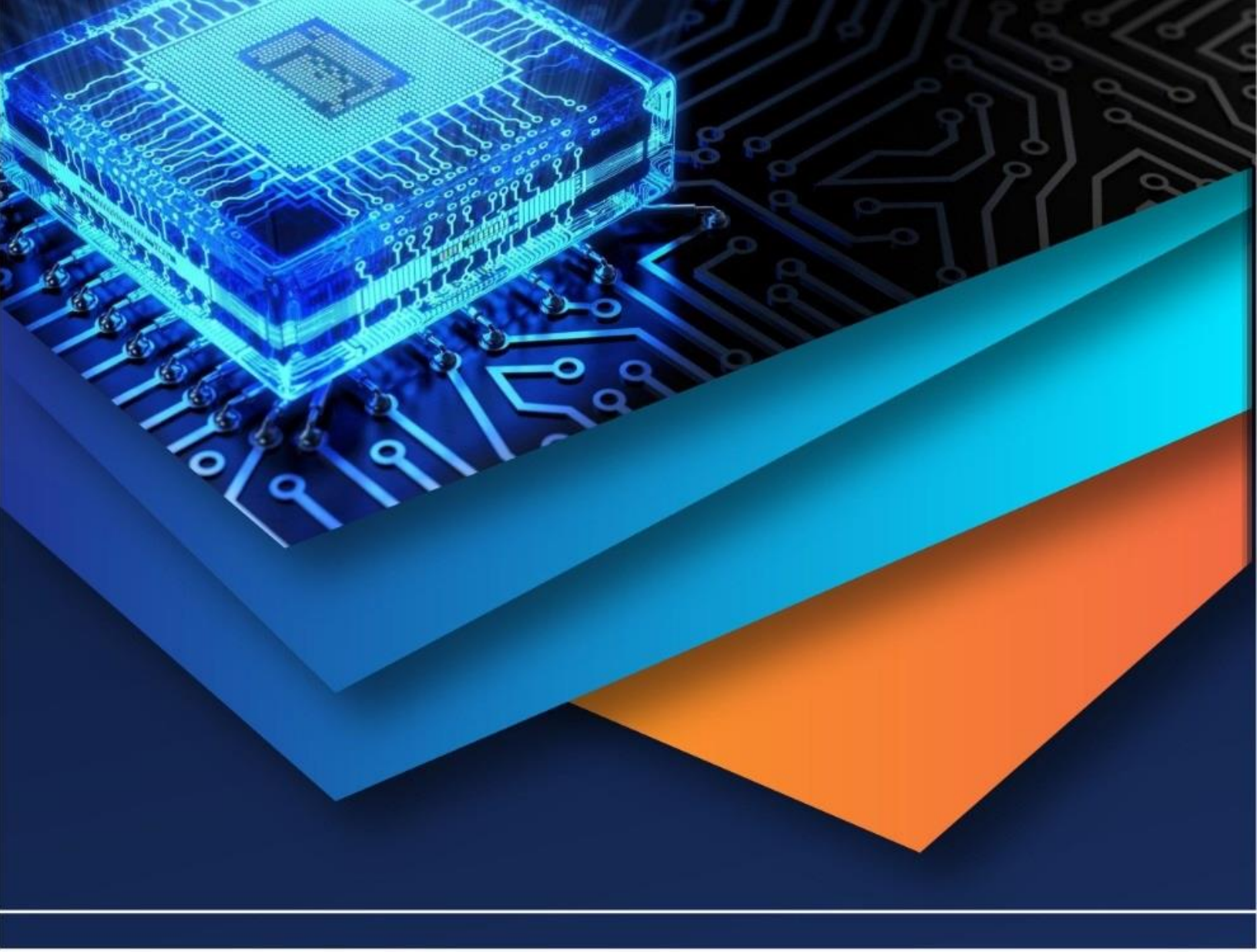
C. Statistical Weather Prediction

Statistical Weather Prediction is used along with the numerical methods. This method is based on past records of weather data and looks for factors that are good indicators for future events. The main purpose is to find out those aspects of weather that are good indicators of the future events. After establishing these relationships, correct data can be safely used to predict the future conditions. Only the overall weather can be predicted in this way. The variables defining weather conditions like temperature (maximum or minimum), relative humidity, rainfall etc., vary continuously with time, forming time series of each parameter and can be used to develop a forecasting model either statistically or using some other means like Artificial Neural Network (ANN) and Decision Tree algorithms.

RFERENCES

[1] Dirmeyer, Paul A.; Schlosser, C. Adam; Brubaker, Kaye L. (February 1, 2009). "Precipitation, Recycling, and Land Memory: An Integrated Analysis". Journal of Hydrometeorology. **10** (1): 278–288.
[2] Fostering Innovation, Creating Jobs, Driving Better Decisions: The Value of Government Data. Economics and Statistics Administration Office of the Chief Economist. July 2014. p. 15. Archived from the original on August 29, 2018. Retrieved December 30, 2018.
[3] D Ramesh, B Vishnu Vardhan, —Crop Yield Prediction Using Weight BasedClustering Technique —, IJCEA, 2015.
[4] Beniwal, S. &Arora, J. (2012), —Classification and feature selection techniques in data mining‖, International Journal of Engineering Research & Technology (IJERT), 1(6)
[5] Xu, R &Wunsch, D (2005), —Survey of clustering algorithms‖, Neural Networks, IEEE Transactions on, 16(3), 645-678.

[6] 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
[7] Rohit Kumar Yadav, Ravi Khatri, "A Weather Forecasting Model using the Data Mining Technique", International Journal of Computer Applications (0975 –8887)Volume 139 – No.14, April 2016
[8] Data mining Models and Algorithms, http://www.huaat.com/english/datamining/D_App.html.
[9] [22]. C.J.C. Burges, *A Tutorial on Support Vector Machines for Pattern Recognition*, Data Mining and Knowledge Discovery 2 (2), 955–974, 1998
[10] Jadhawar BAAM. Weather forecast prediction: A data mining application. International Journal Of Engineering Research And General Science. 2015 Mar-Apr; 4(3):19–21
[11] Chauhan D, Thakur J. Data mining techniques for weather prediction: A review. International Journal on Recent and Innovation Trends in Computing and Communication. 2014 Aug; 2(4):2184–9
[12] Panigrahy, A., Chavan, S.S., Patil, R.N. and Vidyapeeth, B. (2016) An Effective Method for Soil moisture Sensing Using Arduino Uno and Interfacing with GSM Sim900. International Journal for Scientific Research & Development (IJSRD)
[13] Vijayalakshimi, K. and Lakshmi, G.V.M. (2016) Real Time Weather Monitoring from Remote Location Using Raspberry Pi. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJARCCE)
[14] A. Geetha and G. M. Nasira, "Data mining for meteorological applications: Decision trees for modeling rainfall prediction," 2014 IEEE Int. Conf. Comput. Intell. Comput. Res. IEEE ICCIC 2014, pp. 0–3, 2015



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)