Rainfall Detection and Analysis using ML Techniques

Mr. Palukuru Muni[1], Sai Sravan Kumar[2]

Department of Computer Science, Sri Venkateswara University, Tirupati

Abstract— *Intention of this paper is to give non-experts easy access to the techniques and approaches used in the field of rainfall prediction. Rainfall prediction is helpful to avoid flood which save lives and properties of humans. Moreover, it helps in managing resources of water. Information of rainfall in prior helps farmers to manage their crops better which result in growth of country's economy. Fluctuation in rainfall timing and its quantity makes rainfall prediction a challenging task for meteorological scientists. In all the services provided by meteorological department, Weather forecasting stands out on top for all the countries across the globe. We can apply many techniques like classification, regression according to the requirements and also, we can calculate the error between the actual and prediction and also the accuracy.*

I. INTRODUCTION

Rainfall forecasting is very important because heavy and irregular rainfall can have many impacts like destruction of crops and farms, damage of property so a better forecasting model is essential for an early warning that can minimize risks to life and property and also managing the agricultural farms in better way. This prediction mainly helps farmers and also water resources can be utilized efficiently. Rainfall prediction is a challenging task and the results should be accurate. There are many hardware devices for predicting rainfall by using the weather conditions like temperature, humidity, pressure. These traditional methods cannot work in an efficient way so by using machine learning techniques we can produce accurate results. We can just do it by having the historical data analysis of rainfall and can predict the rainfall for future seasons. Most of the rainfall system, Now-a-days are unable to find the hidden layer or any non-linear system patterns present in the system. This project finds hidden layer in the system and provides accurate results. Heavy rainfall prediction is a major problem for meteorological department as it is closely associated with the economy and life of human. It is a cause for natural disasters like flood and drought which are encountered by people across the globe every year. Accuracy of rainfall forecasting has great importance for countries like India whose economy is largely dependent on agriculture. Due to dynamic nature of atmosphere, Statistical techniques fail to provide good accuracy for rainfall recasting. Nonlinearity of rainfall data makes Artificial Neural Network a better technique.

II. LITERATURE SURVEY

M. Fujita, T. Sato, "Observed behaviours of precipitable water vapour and precipitation intensity in response to upper air profiles estimated from surface air temperature", Nature Scientific Reports, vol. 7, no. 1, pp. 4233, 2017.

GPS meteorology is a boom cross-discipline with multiply discipline, and it's a new technology by collecting GPS measuring data to extract meteorological parameters contained in it. Deriving precipitable water vapor from ground-based GPS is one of the important researches. Ground-based GPS network in Chengdu plain is the first network of being continuous and real time in Sichuan region. Chengdu station is the center of the network, other five stations outside constructs a steady polygon measuring frame. This network has run since September 2007. The time series of real time GPS-PWV is retrieved from observation data of being high spatial- temporal resolution, continuous, whole day and real time supplied by GPS network. The application system of water vapor monitoring by ground-based GPS is constructed based on that. It supplies instruction for monitoring atrocious weather and shorttime weather forecast, establishes the foundation for studying spatial-temporal distribution of atmospheric water vapor and exploiting atmospheric water resource through many disciplines.

S. Manandhar, Y. H. Lee, Y. S. Meng, J. T. Ong, "A simplified model for the retrieval of precipitable water vapor from GPS signal", IEEE Trans. Geosci. Remote Sens., vol. 55, no. 11, pp. 6245-6253, Nov 2017.

In this paper, a simplified latitude and day-of-year (DoY)-based model is proposed for the retrieval of precipitable water vapor (PWV) from global positioning system (GPS) signal. Conventionally, PWV, the total amount of water in a vertical column of a unit cross-sectional area, is estimated from the GPS signal delay and a dimensionless conversion factor PI. This PI value is found to rely on a water vapor weighted mean temperature (T_m) value which varies widely across the day, month, and year for different regions. It is, therefore, both time specific and site specific. Analysis of the PI value and its effect on the retrieved PWV from the data obtained for tropical, subtropical, and temperate regions show that although the PI value is time and site specific, the change in the median value of PI for different years is minimal and is dependent only on factors like the latitude coordinates of the particular site and the DoY. Therefore, using the data obtained from 174 different sites, a latitudecoordinate and DoY-based PI value model for the retrieval of PWV is proposed in this paper. The proposed model has been successfully validated using data from different databases: the International GNSS. Service Global Positioning System National Aeronautics and Space Administration (IGS GPS NASA) database, the International GNSS Service Global Positioning System Global Geodetic Observing System (IGS GPS GGOS) database, and the very-long-baseline interferometry (VLBI) database. Results show strong agreement between PWV values calculated using the proposed model and those calculated using the temperature dependent models with 99%, 98%, and 93% of error within ±1 mm for IGS GPS NASA, IGS GPS GGOS, and VLBI databases, respectively. Moreover, the proposed model allows for the ease of PWV retrieval, which is useful in meteorological studies and also applicable in satellite communications.

S. Manandhar, Y. H. Lee, S. Dev, "GPS derived PWV for rainfall monitoring", Proc. International Geoscience and Remote Sensing Symposium (IGARSS), 2016.

Precipitable Water Vapor (PWV) is a good source to monitor precipitation. It is defined by the amount of water vapor present in atmosphere. Traditionally, radiosondes and microwave radiometers were used to derive PWV. However, these devices have poor temporal resolutions and high operational costs. Therefore, GPS signal delay is now widely used for such purposes. The main aim of this paper is to study relationship between GPS derived PWV and precipitation. We present an analysis which shows that PWV increases before any rainfall event, while it decreases after the rainfall event. We also derive a threshold PWV that detects the occurrence of rainfall, once PWV exceeds the threshold value. PWV and rainfall data of June 2010 and 2011 are used for validation.

S. Dev, F. M. Savoy, Y. H. Lee, S. Winkler, "WAHRSIS: A low-cost high-resolution whole sky imager with near-infrared capabilities", Proc. IS&T/SPIE Infrared Imaging Systems, 2014.

Ground-based whole sky imagers are popular for monitoring cloud formations, which is necessary for various applications. We present two new Wide Angle High-Resolution Sky Imaging System (WAHRSIS) models, which were designed specially to withstand the hot and humid climate of Singapore. The first uses a fully sealed casing, whose interior temperature is regulated using a Peltier cooler. The second features a double roof design with ventilation grids on the sides, allowing the outside air to flow through the device. Measurements of temperature inside these two devices show their ability to operate in Singapore weather conditions. Unlike our original WAHRSIS model, neither uses a mechanical sun blocker to prevent the direct sunlight from reaching the camera; instead, they rely on high-dynamic-range imaging (HDRI) techniques to reduce the glare from the sun.

S. Dev, F. M. Savoy, Y. H. Lee, S. Winkler, "Design of low-cost compact and weatherproof whole sky imagers for High-Dynamic-Range captures", Proc. International Geoscience and Remote Sensing Symposium (IGARSS), pp. 5359-5362, 2015

Ground-based whole sky imagers are popular for monitoring cloud formations, which is necessary for various applications. We present two new Wide Angle High-Resolution Sky Imaging System (WAHRSIS) models, which were designed specially to withstand the hot and humid climate of Singapore. The first uses a fully sealed casing, whose interior temperature is regulated using a Peltier cooler. The second features a double roof design with ventilation grids on the sides, allowing the outside air to flow through the device. Measurements of temperature inside these two devices show their ability to operate in Singapore weather conditions. Unlike our original WAHRSIS model, neither uses a mechanical sun blocker to prevent the direct sunlight from reaching the camera; instead, they rely on high-dynamic-range imaging (HDRI) techniques to reduce the glare from the sun.

S. Dev, S. Manandhar, Y. H. Lee, S. Winkler, "Detecting rainfall onset using sky images", Proc. TENCON 2016 – 2016 IEEE Region 10 Conference, 2016.

Ground-based sky cameras (popularly known as Whole Sky Imagers) are increasingly used now-a-days for continuous monitoring of the atmosphere. These imagers have higher temporal and spatial resolutions compared to conventional satellite images. In this paper, we use ground- based sky cameras to detect the onset of rainfall. These images contain additional information about cloud coverage and movement and are therefore useful for accurate rainfall nowcast. We validate our results using rain gauge measurement recordings and achieve an accuracy of 89% for correct detection of rainfall onset.

S. Dev, B. Wen, Y. H. Lee, S. Winkler, "Ground-based image analysis: A tutorial on machine-learning techniques and applications", IEEE Geoscience and Remote Sensing Magazine, vol. 4, no. 2, pp. 79-93, June 2016.

Sky/cloud imaging using ground-based Whole Sky Imagers (WSI) is a cost-effective means to understanding cloud cover and weather patterns. The accurate segmentation of clouds in these images is a challenging task, as clouds do not possess any clear structure. Several algorithms using different color models have been proposed in the literature. This paper presents a systematic approach for the selection of color spaces and components for optimal segmentation of sky/cloud images. Using mainly principal component analysis (PCA) and fuzzy clustering for evaluation, we identify the most suitable color components for this task.

P. K. Wang, Physics and dynamics of Clouds and Precipitation, Cambridge University Press, 2013.

Precipitation retrievals exploiting the available passive microwave (PMW) observations by cross-track and conically scanning satellite-borne radiometers in the Global Precipitation Measurement (GPM) mission era are used to monitor and characterize heavy precipitation events that occurred during the Fall 2014 in Italy. Different physically based PMW precipitation retrieval algorithms are used: the Cloud Dynamics and Radiation Database (CDRD) and Passive microwave Neural network Precipitation Retrieval (PNPR), used operationally in the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) Satellite Application Facility on support to Operational Hydrology and Water Management (H-SAF), and the National Aeronautics and Space Administration (NASA) Goddard Profiling algorithm (GPROF). Results show that PMW precipitation retrievals from the GPM constellation of radiometers provide a reliable and quantitative description of the precipitation (instantaneous and on the daily scale) throughout the evolution of the precipitation systems in the Mediterranean region. The comparable relative errors among gauges, radar, and combination of radiometer overpasses legitimize the use of PMW estimates as a valuable and independent tool for monitoring precipitation. The pixel-based comparison with dual- polarization radars and rain gauges indicate the ability of the different sensors to identify different precipitation areas and regimes (0.60 <; POD <; 0.76; 0.28 <; FAR <; 0.45; 0.42 <; ETS <; 0.59;-1.6 mm/h <; ME <; 1.1 mm/h}, with values depending on the radiometer and on the precipitation product). This is particularly relevant in the presence of complex orography in proximity of coastal areas, as for the analyzed cases. The different characteristics of the radiometers (i.e., viewing geometry, spatial resolution, channel assortment) and of retrieval techniques, as well as the limitations of the ground-based reference datasets, are taken into consideration in ...

P. Benevides, Catalao, P. M. A. Mirandal, "On the inclusion of GPS precipitable water vapour in the nowcasting of rainfall", Natural Hazards and Earth System Sciences European Geosciences Union, vol. 15, pp. 2605-2616, 2015.

GPS-based precipitable water vapor (PWV) estimation has been proven as a cost-effective approach for numerical weather prediction. Most previous efforts focus on the performance evaluation of post-processed GPS-derived PWV estimates using International GNSS Service (IGS) satellite products with at least 3-9-h latency. However, the suggested timeliness for meteorological nowcasting is 5-30 min. Therefore, the latency has limited the GPS-based PWV estimation in real-time meteorological now casting. The limitation has been overcome since April 2013 when IGS released real-time GPS orbit and clock products. This becomes the focus of this paper, which investigates real-time GPS precise point positioning (PPP)-based PWV estimation and its potential for rainfall monitoring and forecasting. This paper first evaluates the accuracy of IGS CLK90 real-time orbit and clock products. Root-meansquare (RMS) errors of <; 5 cm and ~0.6 ns are revealed for real-time orbit and clock products, respectively, during July 4-10, 2013. Second, the real-time GPS PPP-derived PWV values obtained at IGS station WUHN are compared with the post-processed counterparts. The RMS difference of 2.4 mm has been identified with a correlation coefficient of 0.99. Third, two case studies, including a severe rainfall event and a series of moderate rainfall events, have been presented. The agreement between the real-time GPS PPP-derived PWV and ground rainfall records indicates the feasibility of real-time GPS PPP-derived PWV for rainfall monitoring. Moreover, the significantly reduced latency demonstrates a promising perspective of real-time GPS PPP- based PWV estimation as an enhancement to existing forecasting systems for rainfall forecasting.

S. Dev, F. M. Savoy, Y. H. Lee, S. Winkler, "Design of low-cost compact and weatherproof whole sky imagers for High-Dynamic-Range captures", Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 5359-5362, 2015.

Ground-based whole sky imagers are popular for monitoring cloud formations, which is necessary for various applications. We present two new Wide Angle High-Resolution Sky Imaging System (WAHRSIS) models, which were designed specially to withstand the hot and humid climate of Singapore. The first uses a fully sealed casing, whose interior temperature is regulated using a Peltier cooler. The second features a double roof design with ventilation grids on the sides, allowing the outside air to flow through the device. Measurements of temperature inside these two devices show their ability to operate in Singapore weather conditions. Unlike our original WAHRSIS model, neither uses a mechanical sun blocker to prevent the direct sunlight from reaching the camera; instead, they rely on high-dynamic-range imaging (HDRI).

Proposed system:

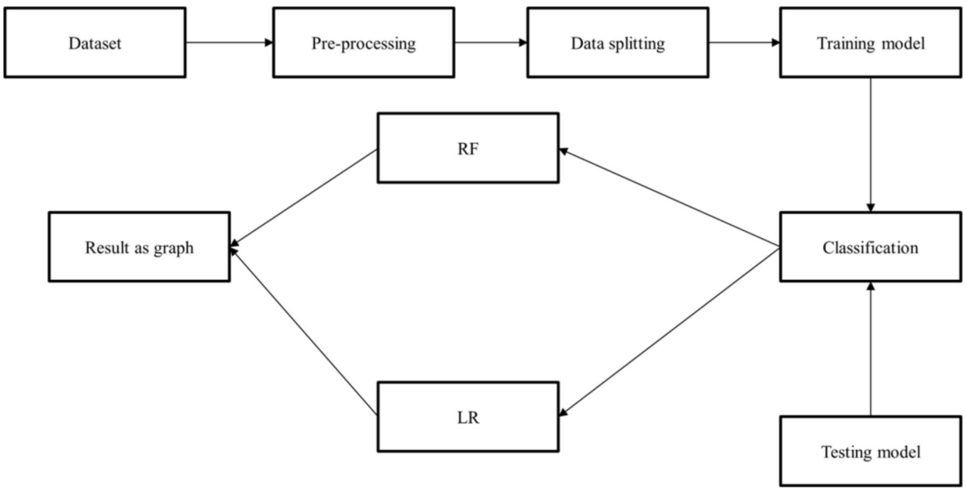
In our Proposed method we are introduce the machine learning models for overcome existing system problems. First, we are collecting all the past experience for the prediction purpose. For the collected dataset divided into two parts. 70 percent of data for training, 30 percent of data for testing. After that implement the RF and LR algorithms. Both algorithms will generate the accuracy score. And also, we are getting some visualization part here based on dataset. The predictive model is used to prediction of the precipitation. The first step is converting data in to the correct format to conduct experiments then make a good analysis of data and observe variation in the patterns of rainfall. We predict the rainfall by separating the dataset into training set and testing set then we apply different machine learning approaches (Random Forest, Logistic Regression, etc.) and statistical techniques and compare and draw analysis over various approaches used. With the help of numerous approaches, we attempt to minimize the error. In proposed system predicts rainfall for the approach which is more accurate. The dataset is collected. There are two techniques to predict rainfall. The first one is logistic regression and the second one is random forest algorithm. This system first compares both the process and then accordingly gives result with the best algorithm. Steps associated with the proposed system are input of data, preprocess of data, splitting of data, training of the algorithm, testing of the dataset, comparing both the algorithms, giving the best algorithm, prediction with the more accurate algorithm and results at the end.

Advantages:

- Improving the accuracy score.
- Reducing the training period for algorithm learning process.
- Increase the dataset size and collect a greater number of features.

Various models and techniques are available today for effective rainfall prediction but still there was a lack of a compact literature review and systematic mapping study which could reflect the current problems, proposed solutions.

III. METHODOLOGY



3.1 Data Collection:

In this module we are collecting past data. This dataset will have different features. Such as sunshine, minimum and maximum temp, etc.

3.2 Preprocessing:

This module we are going to change our dataset into structured format. Here exactly we are implemented pre-processing for removing the NAN or Null values.

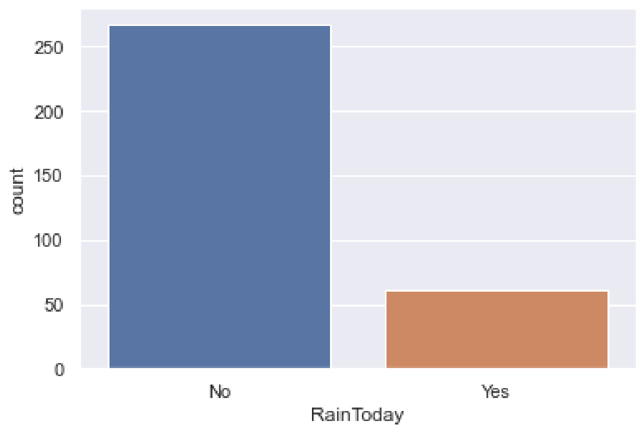
3.3 Data Splitting:

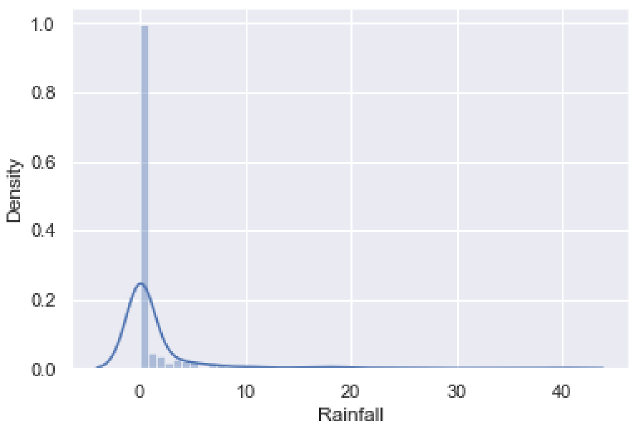
In this module we are going to split our dataset. For the training purpose and testing purpose. Training and testing ratio was 70:30.

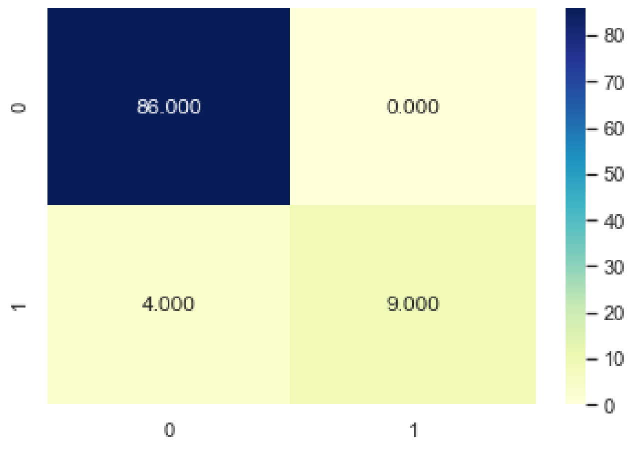
3.4 Prediction:

In this module we need to predict the values based on machine learning models. Here we are going to implement RF and LR.

IV. IMPLEMENTATION

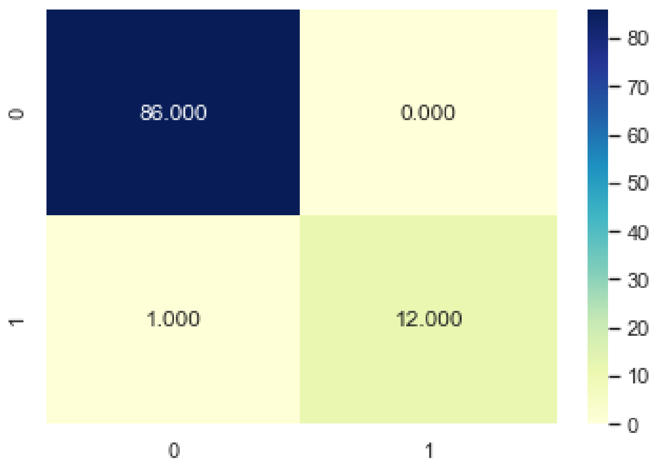


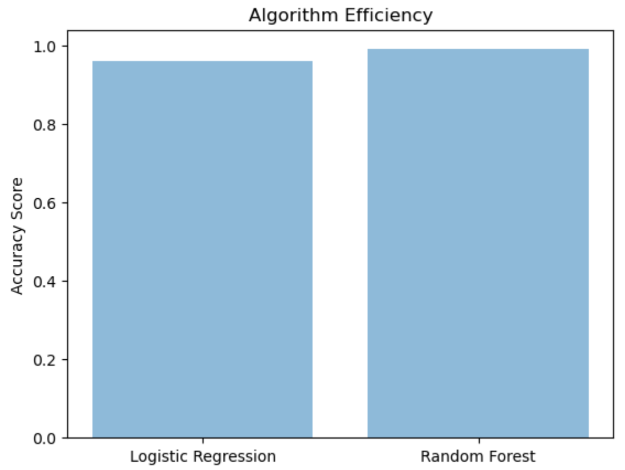




Logistic Regression

Random Forest





REFERENCES

[1] G. A. Vicente, R. A. Scofield, and W. P. Menzel, "The operational GOES infrared rainfall estimation technique," Bull. Amer. Meteorol. Soc., vol. 79, no. 9, pp. 1883–1898, 1998.

[2] R. Joyce, J. Janowiak, and G. Huffman, "Latitudinally and seasonally dependent zenith-angle corrections for geostationary satellite IR bright-ness temperatures," J. Appl. Meteorol., vol. 40, no. 40, pp. 689–703, 2001

[3] A. Y. Hou et al., "The global precipitation measurement mission," Bull. Amer. Meteorol. Soc., vol. 95, pp. 701–722, May 2014.

[4] R. J. Kuligowski, "GOES-R Advanced Baseline Imager (ABI) algorithm theoretical basis document for rainfall rate (QPE), version 2.0," Algorithm Theor. Basis Document (ATBD), Tech. Rep., 2010, pp. 1–44. [Online]. Available: https://www.goesr.gov/products/ATBDs/basele/Hydro_RRQPE_v2.0_no_color.pdf

[5] M. Min, P. Wang, J. R. Campbell, X. Zong, and Y. Li, "Midlatitude cirrus cloud radiative forcing over China," J. Geophys. Res. Atmos., vol. 115, no.D20, p. D20210, 2010.

[6] R. Yu, T. Zhou, A. Xiong, Y. Zhu, and J. Li, "Diurnal variations of summer precipitation over contiguous China," Geophys. Res. Lett., vol. 34, no. 1, pp. 223– 234, 2007.

[7] M. Min, P. Wang, J. R. Campbell, X. Zong, and J. Xia, "Cirrus cloud macrophysical and optical properties over North China from CALIOP measurements," Adv. Atmos. Sci., vol. 28, no. 3, pp. 653–664, 2011.

[8] J. Guo et al., "Delaying precipitation and lightning by air pollution over the Pearl River Delta. Part I: Observational analyses," J. Geophys. Res. Atmos., vol. 121, no. 11, pp. 6472–6488, 2016.

[9] J. Guo et al., "Aerosol-induced changes in the vertical structure of precipitation: A perspective of TRMM precipitation radar," Atmos. Chem. Phys., vol. 18, no. 18, pp. 13329– 13343, 2018.