

Exploratory Analysis of Rainfall Data in India for Agriculture

Team ID

PNT2022TMID10058

Faculty Mentor

Mr. RAJAGOPAL.T. K. P

Team Leader

SRIRAJA B (720719104157)

Team Members

SURESH RAJAN M (720719104160)

THAVAPRAKASH S (720719104167)

SURENDRA GUPTA (720719104153)

Literature Survey

[1] Shikha Srivastava , Nishchay Anand , Sumit Sharma , Sunil Dhar , Lokesh Kumar Sinha, **“Monthly Rainfall Prediction Using Various Machine Learning Algorithms for Early Warning of Landslide Occurrence”** in 2020 International Conference for Emerging Technology (INCET) Belgaum, India

Landslides are considered to be calamitous natural hazards commonly recurring in the Indian Himalayas. Majority of landslides are induced by prolonged or heavy rainfall. Rainfall forecasting helps in identifying the precipitation conditions responsible for landslide occurrence. The proposed research work provides the performance comparison of various machine learning algorithms such as linear regression, back propagation neural network (BPNN), support vector regression (SVR) and long short term memory network (LSTM) used to forecast rainfall which can be compared with the rainfall thresholds to predict landslide occurrence. The analysis is performed using antecedent rainfall data obtained from Narendra Nagar, a small town in the Tehri Garhwal district of Uttarakhand for the period of 1901-2015. Owing to the limited predictability of instantaneous state of the weather, daily rainfall observations are aggregated into monthly indexes. The proposed algorithms use preprocessing techniques followed by data normalization to increase the accuracy of forecasting models. The developed models have the ability to predict rainfall intensity one month in advance or for a specific month of the upcoming year depending upon the dataset used. The study concludes that the BPNNs are able to outperform and provide optimal inferences stating the aptness of artificial neural networks (ANNs) in estimating rainfall and hence predicting the possibility of landslide occurrence well in advance. The study is conducted explicitly for regions highly vulnerable to landslides near Narendra Nagar but may be implemented to any landslide prone area.

[2] T. Dananjali , S. Wijesinghe, J. Ekanayake, **“Forecasting Weekly Rainfall Using Data Mining Technologies”** in 2020 International Conference for Emerging Technology (INCET) Belgaum, India

Rainfall forecasting is a technologically and scientifically a challenging task around the world. Rainfall is one of the most important weather conditions in a given area. Forecasting possible rainfall can help to solve several problems related to the tourism industry, natural disaster management, agricultural industry etc. As the Sri Lankan rural economy is mostly based on agriculture, it is important to forecast rainfall as well as other weather conditions accurately. The weather patterns are localized and hence, generalization of weather prediction models is very difficult. Therefore, this project proposes three data mining models to forecast rainfall, and compares the prediction performances of those models. To that end the data mining models

linear regression, SMO regression, and M5P model were trained from rainfall data collected from the Badulla district, Sri Lanka, during the period 2002 to 2017, to forecast weekly rainfall for the following five months leadtime. Each model was evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Root Relative Squared Error (RRSE), Root Absolute Error (RAE), Direction Accuracy (DA) and residual analysis. According to the findings, the M5P model tree provided the lowest error value, highest direction accuracy, highest correlation between actual and predicted rainfall values, and better randomness of the error values compared to the linear and SMO regression models.

[3] Ashwini, U., Kalaivani, K., Ulagapriya, K., & Saritha, A. (2021),”**Time Series Analysis based Tamilnadu Monsoon Rainfall Prediction using Seasonal ARIMA** on 2021 6th International Conference on Inventive Computation Technologies (ICICT).

Amount of Rainfall prediction is a major issue for the weather department as it is associated with the human's life and the economy. Excess rainfall is the major cause of natural disasters such as drought and flood which are encountered by the people every year across the world. The time series machine learning model is used for forecasting rainfall at Tamilnadu. Forecasting data required for the analysis is available in the Indian meteorological department. To model the monthly rainfall in Tamilnadu for the period from January 1990 to December 2017, the seasonal ARIMA (Auto Regressive Integrated Moving Average) technique is applied. Using the SARIMA (Seasonal Auto Regressive Integrated Moving Average), the stationarity of the time series flow was demonstrated by the rainfall prediction model and then seasonal correlogram assessed. In relation to the Mean Squared Error (MSE) and Root Mean Squared Error, the output of this model is assessed (RMSE). Therefore, it reveals that the ARIMA model accurately forecasts the Rainfall with less error and the derived model could be used to forecast Monsoon rainfall for the upcoming years.