JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Rainfall Prediction Using Lasso Regression Technique

Mr. Shaik Subhan (M.C.A). Rajeev Gandhi Memorial college Of Engineering and Technology, Nandyal
Mr. V.RaviKanth M.Tech,(Ph.D). Rajeev Gandhi Memorial college Of Engineering and Technology, Nandyal

Abstract

Rainfall prediction is the one of the important techniques to predict the climatic conditions in any country. This paper proposes a rainfall prediction model using Linear Regression (LR), Ridge regression, Lasso regression, Decision Tree Regression, Random Forest regression for Indian dataset. The input data is having multiple meteorological parameters and to predict the rainfall in more precise. The Mean Square Error (MSE), accuracy, correlation are the parameters used to validate the proposed model. From the results, the proposed machine learning model provides better results than the other algorithms in the literature.

1. INTRODUCTION

1.1 Introduction

Rainfall prediction is important in Indian civilization and it plays major role in human life to a great extent. It is demanding responsibility of meteorological department to predict the frequency of rainfall with uncertainty. It is complicated to predict the rainfall accurately with changing climatic conditions. It is challenging to forecast the rainfall for both summer and rainy seasons. Researchers in all over the world have developed various models to predict the rain fall mostly using random numbers and they are like the climate data. The proposed model is developed using multiple linear regression. The proposed method uses Indian meteorological date to predict the rain fall.

Usually, machine learning algorithms are classified into two major categories: (i) unsupervised learning (ii) supervised learning. All the clustering algorithms come under supervised machine learning. Figure 1 represents the different classification of machine learning algorithms. Figure 2 describes the rainfall prediction research based on neural network for Indian scenario. Even though many models have developed, but it is necessary for doing research using machine learning algorithms to get accurate prediction. The error free prediction provides better planning in the agriculture and other industries.

2. Literature Survey

• Thirumalai, Chandrasegar, et al. discusses the amount of rainfall in past years according to the crop seasons and predicts the rainfall for future years. The crop seasons are Rabi, Kharif and Zaid. Linear regression method is applied for early prediction. Here, Rabi and kharif were taken as variables if one variable was given then other can be predicted using linear regression. Standard deviation and Mean was also calculated for prediction of crop seasons. This

implementation will be used for farmers to have an idea of which crop to harvest according to crop seasons. Geetha, A., and G. M. Nasira. implements a model which predicts the weather conditions like rainfall, fog, thunderstorms and cyclones which will be helpful to the people to take preventive measures. Data mining techniques were used and a data mining tool named Rapid miner was used to model the decision trees. The data set of Trivandrum with attributes like day, temperature, dew point, pressure etc. The dataset is divided into training set and testing set and decision tree algorithm is applied. The accuracy is calculated, actual and predicted values are compared. The accuracy is 80.67 and to achieve high value it can be extended by applying soft computing techniques like fuzzy logic and genetic algorithms. Parmar, Aakash, Kinjal Mistree, and Mithila Sompura discusses the different methods used for rainfall prediction for weather forecasting with their limitations. Various neural networks algorithm which are used for prediction are discussed with their steps in detail categorizes various approaches and algorithms used for rainfall prediction by various researchers in today's era. Finally, presents conclusion of paper. Done the background work about some models of machine learning ARIMA Model, Artificial neural network and types like Back- Propagation Neural Network - Cascade Forward Back Propagation Network Layer Recurrent Network, Self-Organizing Map and Support Vector Machine, Collected, surveyed and table presents categorization of different approaches of rainfall prediction. Dash, Yajnaseni, Saroj K. Mishra, and Bijaya K. Panigrahi has used artificial intelligence techniques like Artificial Neural Network (ANN), Extreme Learning Machine (ELM), K nearest neighbor (KNN) are applied for prediction of summer monsoon and post monsoon rainfall. The dataset used is the time series data of Kerala from 1871 to 2016 taken from Indian Institute of Tropical Meteorology (IITM).The data is pre-processed and normalization was performed on the data next, the data is divided into training and testing the data up to 2010 was taken as training set and the data from 2011- 2016 taken as test set. The above mentioned algorithms were applied and its performance was calculated by using MAE, RMSE, and MASE. The ELM algorithm has given accurate results compared to the others. Singh, Gurpreet, and Deepak Kumar states that there are many machine learning algorithms applied for the prediction of rainfall and in this, they have used a hybrid approach that is combining two techniques, Random forest and Gradient boosting with many machine learning techniques like ada boost, K-Nearest Neighbor(KNN), Support vector machine(SVM), and Neural Network(NN).These have been applied on the rainfall data of North Carolina from 2007 – 2017 and also the performance is calculated by applying different metrics F-score, precision, accuracy, recall. Finally, eight hybrid models have been proposed and Gradient boosting-Ada boos.

3.1 Existing System:

An existing system is based on the ensemble paradigm include the work in [12], which, similar to our work, employs a probabilistic ensemble and merges two sources of data (i.e., rain gauges and radar) even if the aim of this work is to develop a run-off analysis. Afterward, a blending technique is applied to the results of the runoff hydrologic models to determine a single runoff hydrograph. Experimental results show that the hydrologic models are accurate and can help to make more effective decisions in the flood warning. Frei and Isotta [13] define a technique for deriving a probabilistic spatial analysis of daily precipitation from rain gauges. The final model represents an ensemble of possible fields, conditional on the observations, which can be explained as a Bayesian predictive distribution measuring the uncertainty due to the data sampling from the station network. An evaluation of a real case study, located in the European Alps, proves the capability of the approach in providing accurate predictions for a hydrological partitioning of the region.

3.1.1 Disadvantages of Existing System

- Less feature compatibility
- Low accuracy.

3.2 Proposed System

Our approach is an effective solution for real scenarios, as in the case of an officer of the Department of Civil Protection (DCP), who has to analyze the rainfall in a specific zone presenting risks of landslides or floods. The experimental evaluation is conducted on real data concerning Calabria, a region located in the South of Italy, and provided by the DCP. Calabria is an effective test ground because of its strong climate variability and its complex orography. Our contributions can be summarized as follows. 1) Three heterogeneous data sources (i.e., RGs, radar, and Meteosat) are integrated to generate more accurate estimates of rainfall events. 2) Different classification methods are compared on a real case concerning Calabria, a southern region in Italy, and a hierarchical probabilistic ensemble approach is proposed. 3) Different ML-based methods, pre trained only on historical data, with a widely used interpolation method in the hydrological field (i.e., KED) are compared.

3.3 Methodology

In this project work, I used five modules and each module has own functions, such as:

1. System Module
2. User Module

3.3.1 Dataset Collection:

The dataset to be classified is split into training and testing dataset with the test size of 30-20%.

3.3.2 Preprocessing:

Resizing and reshaping the Dataset into appropriate format to train our model.

.

3.3.3 *Training:*

Use the pre-processed training dataset is used to train our model using Ml algorithm along with some of the Machine learning methods..

3.3.4 Classification

The results of our model is display rain fall prediction.

3.3.5 User Module

Upload Image

The user must upload an dataset which needs to be classified.

View Results

The classified data results are viewed by user.

4 Architecture

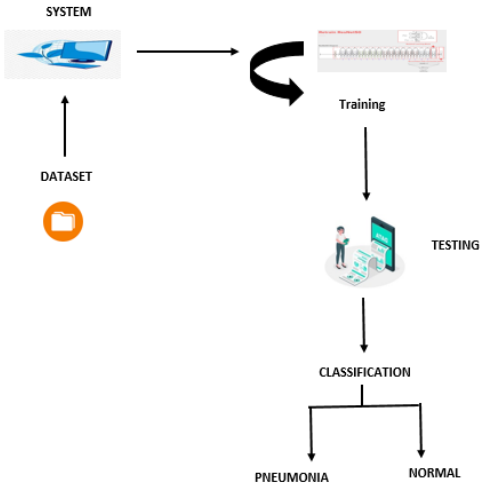
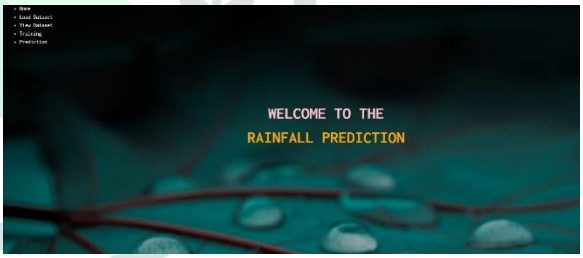


Fig 1: Frame work of proposed method

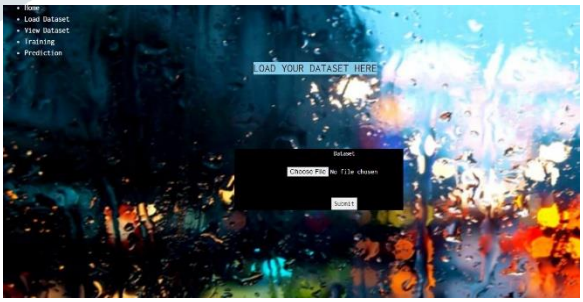
Above architecture diagram shows three stages of data flow form one module to another module. Data collection, preprocessing, and algorithm training.

5 RESULTS SCREEN SHOTS

Home Page:



Upload Data:



Choose options:

Predict Result:

7. CONCLUSION

Rain fall prediction plays the major role in agriculture production. The growth of the agricultural products is based on the rainfall amount. So it is necessary to predict the rainfall of a season to assist farmers in agriculture. The proposed method predicts the rainfall for the Indian dataset using multiple linear regression and provides improved results in terms of accuracy, MSE and correlation.

Future Enhancement

- ✓ This can be utilized in future to predict the rainfall for the Indian dataset using multiple linear regression and provides improved results in terms of accuracy, MSE and correlation.

8. References

1] Manojit Chattopadhyay, Surajit Chattopadhyay, "Elucidating the role of topological pattern discovery and support vector machine in generating predictive models for Indian summer monsoon rainfall", Theoretical and Applied Climatology, pp. 1-12, July 2015, DOI: 10.1007/s00704-015-1544-5

[2] Kumar Abhishek, Abhay Kumar, Rajeev Ranjan, Sarthak Kumar," A Rainfall Prediction Model using Artificial Neural Network", 2012 IEEE Control and System Graduate Research Colloquium (ICSGRC 2012), pp. 82-87, 2012.

[3] Minghui Qiu, Peilin Zhao, Ke Zhang, Jun Huang, Xing Shi, Xiaoguang Wang, Wei Chu, "A Short-Term Rainfall Prediction Model using Multi-Task Convolutional Neural Networks", IEEE International Conference on Data Mining, pp. 395-400, 2017, DOI 10.1109/ICDM.2017.49.

[4] Aswin S, Geetha P and Vinayakumar R, "Deep Learning Models for the Prediction of Rainfall", International Conference on Communication and Signal Processing, April 3-5, 2018, India, pp. 0657-0661.

[5] Xianggen Gan, Lihong Chen, Dongbao Yang, Guang Liu, "The Research Of Rainfall Prediction Models Based On Matlab Neural Network", Proceedings of IEEE CCIS2011, pp. 45- 48.

[6] Sam Cramer, Michael Kampouridis, Alex A. Freitas and Antonis Alexandridis, "Predicting Rainfall in the Context of Rainfall Derivatives Using Genetic Programming", 2015 IEEE Symposium Series on Computational Intelligence, pp. 711 – 718.

[7] Mohini P. Darji, Vipul K. Dabhi, Harshadkumar B.Prajapati, "Rainfall Forecasting Using Neural Network: A Survey", 2015 International Conference on Advances in Computer Engineering and Applications (ICACEA) IMS Engineering College, Ghaziabad, India, pp.706 – 713

[8] Sandeep Kumar Mohapatra, Anamika Upadhyay, Channabasava Gola, "Rainfall Prediction based on 100 years of Meterological Data",2017 International Conference on Computing and Communication Technologies for smart Nation, pp.162 – 166.

[9] Sankhadeep Chatterjee, Bimal Datta, Soumya Sen, Nilanjan Dey, "Rainfall Prediction using Hybrid Neural Network Approach", 2018 2nd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom), pp. 67 – 72.

.