

# Machine Learning based Rainfall Prediction

R. Kingsy Grace<sup>1</sup>,

Department of Computer Science and Engineering  
Sri Ramakrishna Engineering College  
Coimbatore, India

<sup>1</sup>kingsygrace.r@srec.ac.in,B. Suganya<sup>2</sup>

Department of Computer Science and Engineering  
Sri Ramakrishna Engineering College  
Coimbatore, India

<sup>2</sup>suganya.b@srec.ac.in

**Abstract**— Rainfall prediction is the one of the important technique to predict the climatic conditions in any country. This paper proposes a rainfall prediction model using Multiple Linear Regression (MLR) 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.

**Keywords**— Multiple Linear Regression, rainfall, prediction, machine learning, accuracy

## I. 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 similar to 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.

This paper is organized as follows: Section II discusses the various related methods in the literature, Section III explains the proposed method MLR based Rain Fall Prediction. Results are elaborated in section IV and Section V concludes the paper.

## II. RELATED WORK

There are many works in the literature for the prediction of rain fall. This section discusses some of the work related to our proposed methodology. Kumar Abhishek et al. have proposed a rainfall prediction technique using neural network in [2]. The proposed model in [2] predicts the rainfall of Udupi district from Karnataka state of India. BPNN with

feed forward, layer recurrent and BPNN with cascade feed forward neural networks are experimented. The proposed model takes 70% of the data for training and 30% for testing. The recurrent network shows better accuracy when compared to BPNN. The MSE is high in BPNN. Minghui Qiu et al. have introduced rainfall prediction using short term method because of its challenges in the prediction. To solve this convolutional neural network model was used to predict the short term rainfall by collecting set of weather features from multiple surrounding observations. It was compared with public weather forecast model and proved significantly better.

Aswin et al. in [4] proposed a model to predict the rainfall precipitation by using Deep Learning Architectures (LSTM and ConvNet). LSTM and ConvNet Architectures are used to model and predict the Global monthly average rainfall for 10368 Geographic Locations around the globe for 468 Months. RMSE of the proposed system, LSTM is 2.55 whereas the RMSE of ConvNet is 2.44. By increasing the hidden layers errors can be still reduced.

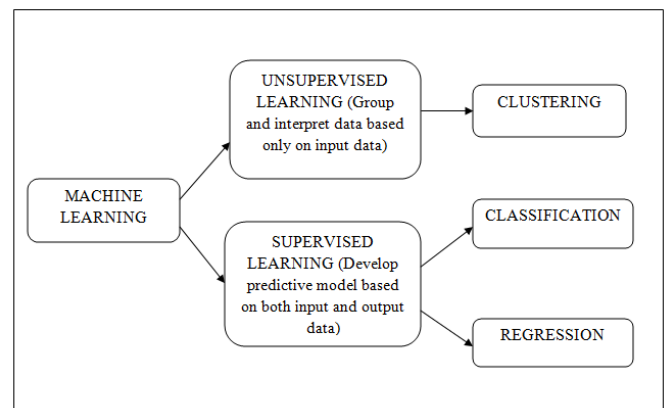


Figure 1. Categorization of Algorithms in Machine Learning

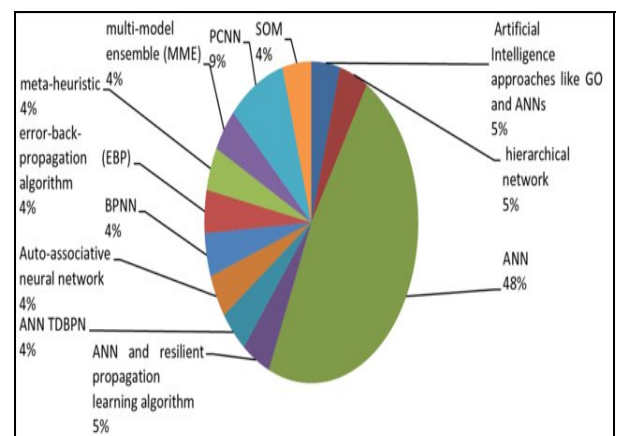


Figure 2. Neural Network based Rainfall Prediction Methodologies [1]

Xianggen Gan et al. have discussed back propagation neural network for rainfall prediction. The Proposed model was tested using the dataset from 1970 to 2000 which has 16 meteorological parameters. During network training the target error is set as 0.01 and learning rate is set as 0.01. Proposed model is implemented using Matlab neural network platform and the accuracy of BPN Network prediction is 100% whereas regression prediction is 67%.

Sam Cramer et al. proposed a methodology to predict rainfall with Genetic Programming. GP and MCRP were compared on 21 different datasets of cities across Europe. Daily rainfall data for 10 years were taken as training data and one year rainfall data were taken as testing data. GP overcomes the weakness of MCRP by predicting the various climates better than MCRP. Mohini P et al. have discussed a survey of different NNs to predict the rainfall. FFNN, RNN, TDNN were used to predict better compared to forecasting techniques. The disadvantage of NNs is that it predicts accurate only for annual rainfall when compared with monthly rainfall prediction. To overcome the problem meteorological parameters can be used. Sandeep Kumar

Mohapatra et al. have proposed rainfall prediction for the data ranging from 1901 to 2002 of Bangalore, India. The Collected data were investigated using data mining technique Linear Regression. To validate and to get computational result pandas and scikit Learn were used. To predict rainfall for different seasons K fold was used. Prediction for rainy season was more accurate compared to summer season.

Sankhadeep Chatterjee et al. have proposed predicting rainfall using Neural Network. The Meteorological Station from Dumdum, West Bengal has given the data in the years between the years 1989 – 1995. The data was clustered using K means clustering. In [9] Hybrid Neural Network (HNN) was compared with MLP-FFN classifier. HNN was better by producing 89.54 % of accuracy (with selection features) and 84.26% of accuracy (without selection features) compared to MLP-FFN.

Sunil Navadia et al. have proposed a model to predict rainfall using predictive analysis in Hadoop. Data were analyzed and prediction of rain was done using Apache PIG. Apache Hadoop can be used in next version to increase the accuracy rate. Table 1 describes the comparison of different rainfall prediction methods in the literature.

### III. MLR BASED RAIN FALL PREDICTION

The proposed method is based on the multiple linear regression. The data for the prediction is collected from the publicly available sources and the 70 percentage of the data is for training and the 30 percentage of the data is for testing. Figure 2 describes the block diagram of the proposed methodology. Multiple regression is used to predict the values with the help of descriptive variables and is a statistical method. It is having a linear relationship between the descriptive variable and the output values. The following is the equation for multiple linear regression:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon$$

Table 1. Comparison of Rainfall Prediction Methods

S. No.	Methods	Performance Parameters	Tools Used
1.	a)Feed Forward with Back –Propagation b)Layer Recurrent c) Cascaded feed Forward back Propagation	Mean Square Error (MSE)	Matlab: Nntool, Nftool
2.	Deep Convolutional Neural Network	MSE, Correlation, Critical Success Index (CSI)	Not Mentioned
3.	LSTM and ConvNet	Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE)	Not Mentioned
4.	BP network	Accuracy	Matlab: Neural Network Platform
5.	Genetic Programming	RMSE	Not Mentioned
6.	Artificial Neural Network	Accuracy	Meteorological Parameters
7.	Linear Regression	RMSE, MAE	Pandas and scikit Learn
8.	Hybrid Neural Network	Accuracy, Precision, Recall	Not Mentioned
9.	Likelihood	Accuracy	Hadoop

The number of observation is indicated by  $n$ . The dependent variable is  $y_i$  and the descriptive variable is  $x_i$ .  $\beta_0$  and  $\beta_p$  are the constant y intercept and slop of descriptive variable respectively. Model error is indicated by  $\varepsilon$ . In the proposed model multiple meteorological parameters are necessary to predict the rain fall, it is better to use the multiple linear regression instead of simple linear regression. The assumptions which are made by the multiple linear regression are: linear relationship between the both the descriptive and independent variables, the highly correlated variables are independent variables,  $y_i$  is calculated randomly and the mean and variance are  $\theta$  and  $\sigma$ . Figure 3 explains the flow of MLR prediction.

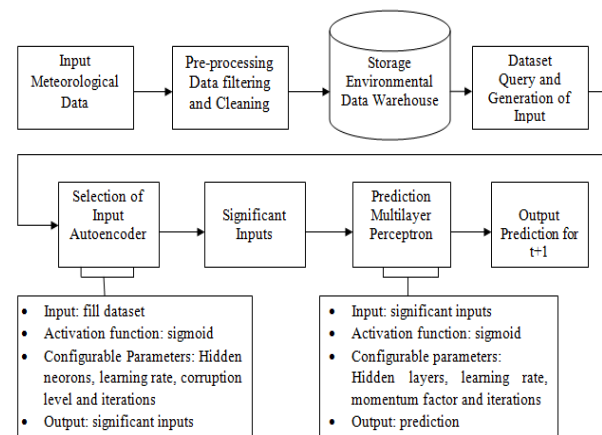


Figure 2. Block Diagram of Proposed Methodology

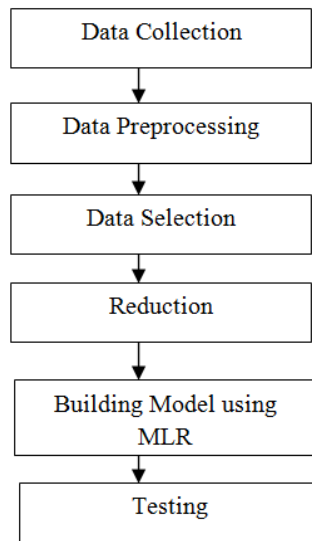


Figure 3. Model Generation using MLR

#### IV. RESULTS

This section deals with the results in the proposed MLR based rain fall prediction method. The total number of data in the selected data set is 4116. Figure 4 describes the MLR prediction result. The accuracy of the MLR prediction is 0.99 and is shown in Figure 5. The comparison of the performance parameters is shown in Table 2.

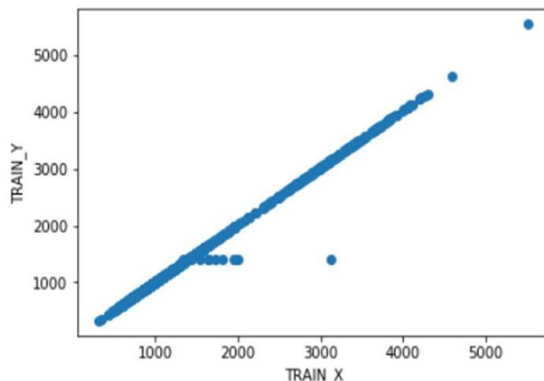


Figure 4. MLR Prediction Result

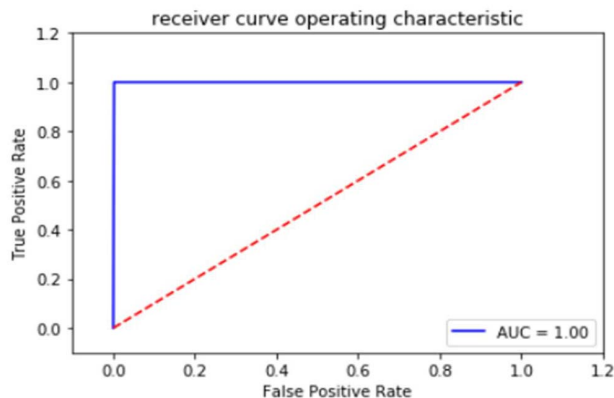


Figure 5. MLR Accuracy after Prediction

Table 2. Comparison of the Performance Parameters

S. No.	Algorithm	MSE	RMSE	Correlation
1.	QPF	15.547	3.943	0.399
2.	LR	13.28	3.644	0.469
3.	MLR	11.894	3.449	0.473

#### V. 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.

#### 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 Meteorological 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.
- [10] Mr. Sunil Navadia, Mr. Pintukumar Yadav, Mr. Jobin Thomas, Ms. Shakila Shaikh, "Weather Prediction: A novel approach for measuring and analyzing weather data", International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2017), pp. 414 - 417