ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 1, 2025

Rainfall Prediction Using Machine Learning: A Comparative Analysis of MLR and Artificial Neural Networks

[1] Mr.M. Rajakannan, Assistant Technical Officer

[2] M.Ramesh, Assistant Professor,

[1] Department of Genetics and plant Breeding, Annamalai University, Cuddalore(Dist), Tamilnadu.
[2] Department of Information Technology, St. Martin's Engineering College, Dhulapally, Secunderabad, Telangana.

ABSTRACT

Rainfall plays a vital role in the economy of countries where agriculture is the primary source of income, such as India. Accurate rainfall forecasting is crucial not only for effective agricultural planning but also for safeguarding communities, particularly those in coastal regions, from risks associated with heavy rain and flooding. Early and reliable predictions allow individuals to take preventive measures, reducing potential losses and improving overall preparedness.

This study focuses on developing an efficient rainfall prediction model using Multiple Linear Regression (MLR) and Artificial Neural Networks (ANN). The research compares the performance of various machine learning algorithms, including MLR, Neural Networks, K-means, and Naïve Bayes, to determine the most suitable approach for rainfall forecasting. The input dataset integrates multiple meteorological parameters such as humidity, temperature (minimum and maximum), pressure, cloud cover, and wind speed to enhance prediction accuracy.

The proposed model is evaluated using metrics like Mean Absolute Error (MAE), accuracy, and correlation. Results indicate that the suggested approach outperforms existing methods in terms of precision and reliability, making it a promising solution for rainfall prediction and its associated applications in agriculture and disaster management.

Keywords: Rainfall prediction, Machine learning algorithms, MLR, Artificial Neural network

1. INTRODUCTION

Rainfall prediction is important in Indian civilization and it plays a major role in human life to a great extent. It is a demanding responsibility of the meteorological department to predict the frequency of rainfall with certainty. It is complicated to predict the rainfall accurately with changing climatic conditions. It is challenging to forecast the rainfall for both the summer and rainy seasons. Researchers all over the world have developed various models to predict the rainfall mostly using random numbers and they are similar to the climate data.

The Rainfall Prediction model is implemented by using two Algorithms which are Multiple Linear Regression and Neural Networks. Rainfall Prediction using Linear Regression and Neural Networks is to find the correlation between diverse features in the dataset which contribute to Rainfall and to find correct weights and Biases which lead to accurate Prediction of Rainfall respectively. Initially, the dataset with multiple featuresis cleaned and pre-processed to make it suitable for use and feed into a machine learning algorithm.

There are many hardware devices for predicting rainfall by using the weather conditions like temperature, humidity, and pressure. These traditional methods cannot work in an efficient way so by using machine learning techniques we can produce accurate results. We will just do it by having the historical data of rainfall and other parameters like temperature, pressure, humidity, cloud, wind can predict the rainfall for future seasons. We can apply many techniques like classification, and regression according to the requirements, and also we can calculate the error between the actual and prediction and also the accuracy. Different techniques produce different accuracies so it is important to choose the right algorithm and model it according to the requirements.

Form the last few decade scientist and engineers are successfully produced several models for making accurate predictions in several fields. Machine learning is also a field that is widely used for prediction purposesor for classifying things. There are a number of methods, listing from KNN, more complex methods such as SVM and ANN (Artificial Neural Network). For metrology predictions, ANNs pictured as an alternative methodas opposed to the traditional method, are based on self-adaptive mechanisms that learn from examples and capture functional relationships between data, even if the relationships between the data are unknown or difficult to describe.

2. LITERATURE SURVEY

According to R. Kingsy Grace et.al [1] one of the most important techniques for predicting meteorological conditions in any country is rainfall prediction. For the Indian dataset, this paper proposes a rainfall prediction model based on Multiple Linear Regression (MLR). Multiple meteorological characteristics are included in the input data, allowing for a more accurate estimate of rainfall. The suggested model is validated using the Mean Square Error (MSE), accuracy, and correlation metrics. The proposed machine learning model outperforms all other techniques in the literature, according to the results.

Deepali Patil et.al [2] says that weather and rainfall, in general, are highly nonlinear and complicated phenomena that necessitate advanced computer modeling and simulation for accurate prediction. To predict rainfall, a variety of machine learning models are utilized, including Multiple Linear Regression, Neuralnetworks, K-means, Nave Bayes, and others. By extracting, training, and testing data sets and identifying and predicting rainfall, these systems accomplish one of these applications. This paper shows the use of Multiple Linear Regression and Neural networks to predict rainfall and the Decision Trees algorithm to recommendcrops. Thus, concluded that we can reasonably anticipate rainfall and propose crops.

According to Aakash Parmar et.al [3] predicting heavy rainfall is a big concern for meteorological departments since it is so directly linked to the economy and human existence. It is the cause of natural disasters such as floods and droughts that affect people all over the world every year. Rainfall forecasting accuracy is critical for countries like India, whose economy is heavily reliant on agriculture. Statistical techniques fail to provide good accuracy for rainfall forecasting due to the dynamic character of the atmosphere. Because of the nonlinearity of rainfall data, Artificial Neural Networks is a superior technique. In a tabular format, the review work and comparison of different methodologies and algorithms employed by researchers for rainfall prediction are displayed. The goal of this paper is to make the techniques and approaches used in this study accessible to non- experts.

According to Urmay Shah et.al [4] estimating precipitation is one of the most important aspects of meteorological science. To forecast and estimate meteorological parameters, a couple of factual procedures and machine learning approaches are used to forecast and estimate precipitation. The purpose of this paper is to provide climatic insights to clients from diverse industries, such as agriculturists, researchers, and others, so thatthey may understand the relevance of changes in climate and atmosphere characteristics such as precipitation, temperature, and humidity.

According to Arnav Garg et.al [5] as global warming worsens, detecting and forecasting rainfall is becoming a serious concern in countries without enough technology, which, if done correctly, can aid them in a variety of ways, including agricultural, health, drinking, and many more. This paper compares the findings inferred by each technique and uses SVR, SVM, and KNN machine learning algorithms to estimate the rainfall for the future year.

According to Kaushik Dutta et.al [6] rain prediction models based primarily on artificial neural networks have been proposed in India. This paper compares two rainfall forecast methods and determines which one is the most accurate. With the complicated data available, the current technique for predicting rainfall does not function effectively. Statistical and numerical methods are now in use, but they do not operate accurately when there is a non-linear trend. When the complexity of the datasets containing historical rainfall rises, the existing approach fails. To determine the best technique to predict rainfall, researchers will look at both machine learning and neural networks, and the algorithm that delivers the greatest results will be employed in prediction. Rainfall estimation is required to build an effective rainwater harvester. Weather forecasting is the simplest and quickest way to get a greater outreach. This research can be used by all-weather forecasting channels, allowing for more accurate forecasting news to reach all sections of the country.

3. METHODOLOGY

3.1. Dataset Description

The dataset consists of the measurement of rainfall from the year 2009 to 2021 for Kolhapur from the website World Weather Online. The data describe meteorological information from Kolhapur District collected daily for 12 years. In particular, some parameters indicate whether it rains on the same day, Rain Today, which will be the target variable that will be tested to predict using machine learning algorithms.

- Data consists of parameters such as Minimum temperature, Maximum temperature, Pressure,

Humidity, Wind, Cloud, and rain.

- 20% of data is used for Testing and 80% of data is used for Training.
- Rain is the dependent parameter and other parameters like Minimum temperature, Maximum temperature, Pressure, Humidity, Wind, and Cloud are independent variables.
- The attributes are the amount of rainfall measured in mm.

3.2. Data Preparation

1. Missing data

If the samples that do not have data in any of their variables are eliminated, then approximately 50% of the samples would have to be eliminated. That is why the variables for which there are no data were analyzed, separating them by cities in order not to discard a large amount of data.

2. Data normalization

Normalization of the data of the mix–max type was carried out so that all the variables would take values between 0 and 1. And like this variables taking values of great magnitudes having a greater influence on the application of machine learning algorithms were avoided.

3. Detection of outliers

The outliers within the graph are removed and so accuracy is checked because due to that outlier's accuracy of the model is reduced and as a result accuracy is improved.

3.3. Building a Model (Multiple Linear Regression and Artificial Neural Networks)

1. Multiple Linear regression

In statistics, linear regression will be defined as a linear approach to demonstrating the correlation between a dependent variable and one or additional independent variables. In this case, if there is just one independent variable, it is called simple linear regression. If there is more than one independent variable, the method is called multiple linear regression. This term features a totally different means from multivariate linear regression, wherever multiple correlated dependent variables are predicted, rather than a single variable. This linear regression model has many applications such as prediction, forecasting, and error reduction. A predictive model will be fitted using Linear regression with the collected dataset or discovered values. Once the model is fitted, it will be used to create predictions. The multiple linear regression equation is as follows: Once the model is fitted, it will be wont to create predictions.

The multiple linear regression equation is as follows:

$$y = a_1 + a_2x_1 + a_3x_2 + a_3x_4 + a_5x_5 + a_6x_6 . .$$

wherever y is the predicted rainfall value, $a_1, a_2, a_3, a_4, a_5, a_6$ are the predictor coefficients, and

$x_1, x_2, x_3, x_4, x_5, x_6$ are the predictor variables.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 1, 2025

2. Artificial Neural Networks

Artificial Neural Networks is one in all most popular machine learning and deep learning algorithms. They are impressed by human neurons that are capable of creating human like decisions with facilitate of computations. For instance, in our case we have tendancy to trained the Neural Networks with totally different features like humidity, temperature, pressure etc. and that they learn to identify and analyze the rainfall supported these features using the results of training dataset. The very simpleneural network would possibly contain only one input neuron, one hidden neuron, and one output neuron. It takes several dependent variables that is equal to input parameters, multiplies them by their coefficients equal to weights, and runs them through a ReLU activation function and a unit step function.

$oj = f(\sum wi, jai + bi)$

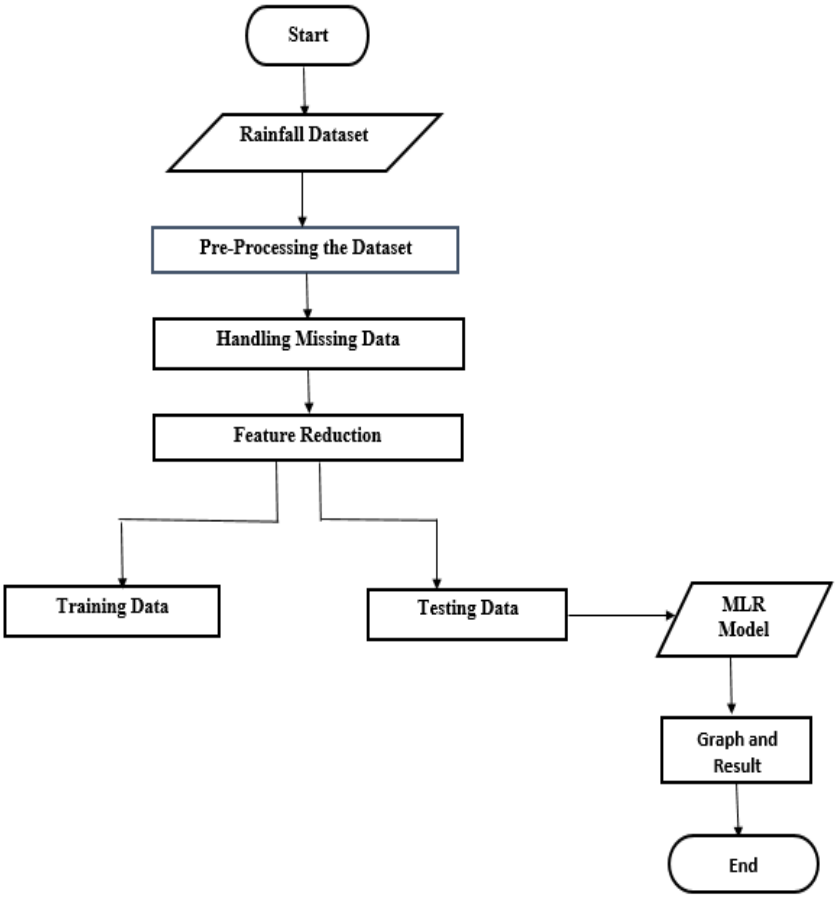

Fig-1: Rainfall Prediction Model

Step 1. Rainfall Dataset: Import the rainfall data set CSV file. The dataset consists of the measurement of rainfall from the years 2011 to 2021 for Kolhapur. Data consists of parameters such as Minimum temperature, Maximum temperature, Pressure, Humidity, Wind, etc. The attributes are the amount of rainfall measured in mm.

Step 2. Pre-processing A preliminary process of data so as to organize it for the primary processing or for any further analysis. The term can be applied to any first or preparatory process stage once there are many steps needed to organize data for the user. Fill the missing values with the average that is mean value of the data.

Step 3. Handling Missing Data: Scaling the features- scaling the data to a fixed scale. The reason of missing values will be data corruption or failure to record data. The handling of missing data is extremely necessary throughout the preprocessing of the dataset as many machine learning algorithms don't support missing values.

Step 4. Feature Reduction: PCA is used to minimize the data. Feature reduction, called as dimensionality reduction, is the method of reducing the number of features in a resource-heavy computation without losing very useful information.

Step 5. Training and Testing: Training and Testing is a method to measure the accuracy of our model. It is called Train/Test because you split the data set into two sets: a training set and a testing set. 80% for training, and 20% for testing. You train the model using the training set. The data is divided into a training set (80%) and a testing set (20%).

Step 6. Prediction Model: Multiple Linear Regression algorithm, Support Vector Regression, and Lasso Regression is applied and the Mean Absolute Error, r2 score is calculated.

Step 7. Graph: The scatter plots are plotted between predicted and testing data for the applied models and the errors are compared and the best model among them is selected.

Step 8. Results: Display the results and gives the desired output with more accuracy.

4. RESULTS AND DISCUSSION

This section deals with the results in the proposed MLR and Neural Network based rain fall prediction method. In this project core motive is to finding out the algorithm which gives us the good prediction of rainfall. Here we took the rainfall data of Kolhapur of past 12 years from 2009-2021 from the official site world weather online. Below is the table of the accuracy of the algorithms.

Sr. No.	Algorithms	MAE
1	Multiple Linear Regression	57
2	Neural network	42

Table-1: Comparative Results

5. CONCLUSION

Rainfall prediction plays a major role in agriculture production. The amount of rainfall determines the growth of agricultural crops. As a result, it is vital to predict seasonal rainfall in order to assist farmers in agriculture. In this project two methods are being examined in order to find an effective approach to predict rainfall. One is the multiple linear regression approach and the second one is the artificial neural networks approach and provides improved results in terms of accuracy, MSE, and correlation.

6. REFERENCES

1. Grace, R. K., & Suganya, B. (2020, March). Machine learning based rainfall prediction. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)* (pp. 227-229). IEEE.
2. Patil, D., Badarpura, S., Jain, A., Student, U. G., & Gupta, A. Rainfall Prediction using Linear approach & Neural Networks and Crop Recommendation based on Decision Tree.
3. Parmar, A., Mistree, K., & Sompura, M. (2017, March). Machine learning techniques for rainfall prediction: A review. In *International Conference on Innovations in information Embedded and Communication Systems* (Vol. 3).
4. Shah, U., Garg, S., Sisodiya, N., Dube, N., & Sharma, S. (2018, December). Rainfall prediction: Accuracy enhancement using machine learning and forecasting techniques. In *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)* (pp. 776-782). IEEE.
5. Garg, A., & Pandey, H. (2019). Rainfall prediction using machine learning. *International Journal of Innovative Science and Research Technology*, *4*(5), 56-58.
6. Dutta, K., & Gouthaman, P. (2020). Rainfall prediction using machine learning and neural network. *International Journal of Recent Technology and Engineering (IJRTE)*, *9*(1), 1954-1961.