

RESEARCH

Open Access

Machine learning techniques to predict daily rainfall amount



Chalachew Muluken Liyew*  and Haileyesus Amsaya Melese

*Correspondence:
chalachewsweet@gmail.com
Bahir Dar University, Bahir Dar
Institute of Technology, Bahir
Dar, Ethiopia

Abstract

Predicting the amount of daily rainfall improves agricultural productivity and secures food and water supply to keep citizens healthy. To predict rainfall, several types of research have been conducted using data mining and machine learning techniques of different countries' environmental datasets. An erratic rainfall distribution in the country affects the agriculture on which the economy of the country depends on. Wise use of rainfall water should be planned and practiced in the country to minimize the problem of the drought and flood occurred in the country. The main objective of this study is to identify the relevant atmospheric features that cause rainfall and predict the intensity of daily rainfall using machine learning techniques. The Pearson correlation technique was used to select relevant environmental variables which were used as an input for the machine learning model. The dataset was collected from the local meteorological office at Bahir Dar City, Ethiopia to measure the performance of three machine learning techniques (Multivariate Linear Regression, Random Forest, and Extreme Gradient Boost). Root mean squared error and Mean absolute Error methods were used to measure the performance of the machine learning model. The result of the study revealed that the Extreme Gradient Boosting machine learning algorithm performed better than others.

Keywords: Machine learning, MLR, RF, XGBoost, Rainfall prediction

Introduction

Based on the distribution of rainfall in Ethiopia, three distinct seasons are identified which are Belg, Kiremt and Bega. According to Ehsan et al. [1] three seasons are; the 'short' rains (belg: February–May), followed by the long rains (kiremt: June–September) and the dry season (Bega: October–January). Kiremt is the main Ethiopian rainy season, and Ethiopia receives a substantial fraction of its annual rainfall during this season, which is very important for its water resources management and agriculture production. The northwestern part of the country at which this research is conducted experiences higher rainfall amounts from June to September that send a flood into the Blue Nile. Droughts and floods have been a major and persistent challenge of the management of water resources, agro-economic, livestock growth, and food production in Ethiopia. To use the rainfall water efficiently, rainfall prediction is unquestionable research area in Ethiopia.

Rainfall prediction is crucial for increasing agricultural productivity which in turn secures food and quality water supply for citizens of one's country. The scarcity of rainfall has a negative influence on the aquatic ecosystem, quality water supply, and agricultural productivity. Agriculture and water quality depend on the rainfall and water amount on a daily and annual basis [2–4]. Therefore, accurate prediction of daily rainfall is a challenging task to manage the rainfall water for agriculture and water supply.

Various researchers conducted studies to improve the prediction of daily, monthly and annual rainfall amounts using different countries' meteorology data. Researchers applied data mining techniques [2, 3, 5, 6] Big Data analysis [4, 7], and different machine learning algorithms [8–11] to improve the accuracy of daily, monthly and annual rainfall prediction. According to the results of the studies, the prediction process is now shifted from data mining techniques to machine learning techniques. Scholars, for example [4], confirmed that machine learning algorithms are proved to be better replacing the traditional deterministic method to predict the weather and rainfall. Consequently, this paper analyzed different machine learning algorithms to identify the better machine learning algorithms for accurate rainfall prediction.

Several environmental factors affect the existence of rainfall and its intensity. The temperature, relative humidity, sunshine, pressure, evaporation, etc. are some of the factors that affect the existence of rainfall and its intensity directly or indirectly. The study conducted by Chaudhari and Choudhari [12] indicated that temperature, wind, and cyclone were important features of the atmosphere over the Indian region to predict rainfall, however, the study did not measure the correlations of each feature to determine the strength of the independent features on the rainfall. On the other hand, a correlation study by Thirumalai et al. [13] identified the most important features like solar radiation, perceptible water vapor, and diurnal features for rainfall prediction using a linear regression model. Whereas, scholars (for example, [10, 11, 14]) used atmospheric features of temperature, relative humidity, pressure, and wind speed as an important feature to predict rainfall accurately using machine learning such as Artificial Neural Network, Random forest, and multiple linear regression model respectively. Hence, important atmospheric features that have a direct or indirect impact on rainfall should be studied to predict the existence and the intensity of rainfall.

Therefore, this study aimed to identify the relevant atmospheric features that cause rainfall and predict the intensity of daily rainfall using machine learning techniques. The raw data is collected from regional meteorology and preprocessed to make it suitable for the experiment. Each feature of the preprocessed data is correlated with the rainfall variable to identify the relevant features using Pearson correlation. The study then experimented the Radnom forest (RF), MLR and XGBoost machine learning algorithms. The MAE and RMSE values of the XGBoost gradient descent algorithms were 3.58 and 7.85 respectively so that The XGBoost algorithm predicted the rainfall using relevant selected environmental features better than the RF and the MLR.

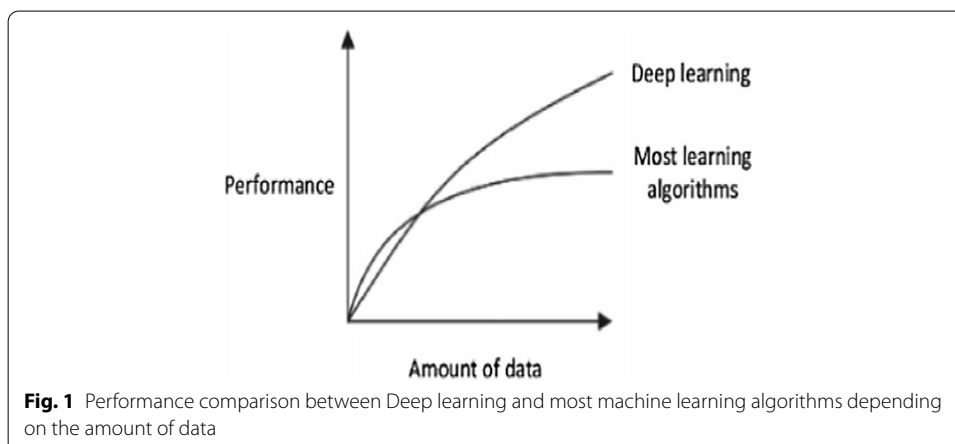
Related work

The machine learning algorithm called linear regression is used for predicting the rainfall using important atmospheric features by describing the relationship between atmospheric variables that affect the rainfall [13, 15]. The correlation study is

conducted [7], and identified solar radiation, perceptible water vapor, and diurnal features are important variables for daily rainfall prediction using a data-driven machine learning algorithm. The future work identified by Manandhar et al. [7] is studying the impact of using different atmospheric features using a larger data set. The researches address the relationship between independent and dependent features to identify which features impact the rainfall to rain or not to rain. The amount of daily rainfall was not found or addressed in this research, it may reduce the performance of the system. Tharun et al. [5] performed the accuracy measure of the comparative study of statistical modeling and regression techniques (SVM, RF & DT) for rainfall prediction using environmental features. According to the result of the study, the regression techniques of rainfall prediction outperformed the statistical modeling. The experimental result showed that the RF model performed and predicted accurately than the SVM and DT. Hence, rainfall prediction is accurate, it shows high performance in machine learning models than the traditional models. This research used different machine learning techniques rather than statistical methods to predict daily rainfall amounts.

The study by Arnav Garg and Kanchipuram [8] shows three machine learning algorithm experiments such as support vector machine (SVM), support vector regression (SVR), and K-nearest neighbor (KNN) using the patterns of rainfall in the year. The SVM algorithm performs best among the three machine learning algorithms. This research did not show the experiment result that which environmental features impact the intensity of rainfall. This paper shows the environmental features that have a positive and negative impact on rainfall and predicts the daily rainfall amount using those features.

Scholars, for example, [14, 16] confirmed that the multiple linear regression machine learning algorithm outperforms well to predict rainfall using dependent weather variables of temperature, humidity, moisture, wind speed, and finally the study showed the performance of the rainfall prediction can be improved using deep learning models as future work. According to Sarker [17, 18] the performance comparison between deep learning and other machine learning algorithms has been shown in Fig. 1 below, where the deep learning model performance increases when the size of the data is increased. Due to the size of the data that is used in this study, machine learning techniques are appropriate.



Scholars [9, 10] studied the deep learning algorithm for rainfall prediction by using different dependent weather variables. To provide an accurate prediction of rainfall, prediction models have been developed and experimented with using machine learning techniques.

Therefore, most researchers did not show the prediction of the daily rainfall amount rather conducting experiments on environmental data to predict whether rain or not rain and predict average annual rainfall amount that is the prediction of daily rainfall amount is a challenging task. All relevant environmental features important for rainfall prediction were not used. This paper examined the machine learning algorithms using data collected from one meteorology station which is relatively small in size and selected the appropriate environmental features that correlate with rainfall positively or negatively to examine the performance of the daily rainfall amount prediction machine learning algorithms using MAE and RMSE.

Machine learning algorithms

To choose the better machine learning algorithms to study the daily rainfall amount prediction, various papers have been reviewed concerning rainfall prediction. To predict the daily rainfall intensity using the real-time environmental data, three algorithms such as MLP, RF, and XGBoost gradient descent were chosen for the experiment. Hence, the three machine learning algorithms were experimented with and compared to report the better algorithms to predict the daily rainfall amount.

Multivariate linear regression (MLR)

Linear regression can be multivariate which has multiple independent variables used as input features and simple linear regression which has only one independent or input feature. Both linear regressions have one dependent variable which can be forecasted or predicted based on the input features. This paper presented the multivariate linear regression because multiple environmental variables or features were used to predict the dependent variable called daily rainfall amount. Linear regression is a supervised machine learning technique used to predict the unknown daily rainfall amount using the known environmental variables. The multivariate linear regression used multiple explanatory or independent variables (X) and single dependent or output variable denoted by Y . Hence, the general equation of the multiple linear regression is given as:

$$Y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \dots + \beta_p x_{ip} + \varepsilon_i = x_i^T \beta + \varepsilon_i \quad i = 1, 2, 3 \dots n$$

where x_i^T is transpose of x_i the input or independent variable, β is regression coefficient, ε_i is error term or noise, Y_i is a dependent variable.

The general multivariate linear regression equation of this paper is given as

$$\begin{aligned} \text{Daily rainfall} = & (\text{year} * \beta_1) + (\text{month} * \beta_2) + (\text{day} * \beta_3) + (\text{MaxTemp} * \beta_4) \\ & + (\text{MinTemp} * \beta_5) + (\text{Humidity} * \beta_6) + (\text{Evaporation} * \beta_7) \\ & + (\text{sunshine} * \beta_8) + (\text{windspeed} * \beta_9) + \varepsilon_i \end{aligned}$$

The size of the data set collected from the meteorological station for this study was appropriate to use the machine learning algorithms called multivariate linear regression that can estimate the daily amount of rainfall in the region. This algorithm can show how strongly each environmental variable influences the intensity of the daily rainfall.

Random forest (RF)

A Random Forest Regression model is powerful and accurate. It usually performs great on many problems, including features with non-linear relationships. Random forest regression is a supervised machine learning algorithm that uses the ensemble learning method for regression. RF works by building several decision trees during training time and outputting the mean of the classes as the prediction of all the trees. The RF algorithm works on the following steps:

- a. Take at random p data points from the training set
- b. Build a decision tree associated with these p data points
- c. Take the number N of trees to build and repeat a and b steps
- d. For a new data point, make each one of the N tree trees predict the value of y for the data point and assign the new data point to the average of all of the predicted y values.

Random forest algorithm is one of the supervised machine learning algorithms that are selected as the predictive model for daily rainfall prediction using environmental input variables or features. Random forest regression is operated by constructing a multitude of decision trees at the training time and outputting the class that is the mode of mean prediction or regression of the individual trees. According to [2] the RF algorithm is efficient for large datasets and a good experimental result is obtained using large datasets having a large proportion of the data is missing.

XGBoost gradient descent

XGBoost stands for eXtreme Gradient Boosting; it is a specific implementation of the Gradient Boosting method which uses more accurate approximations to find the best tree model. XGBoost is implemented for the supervised machine learning problem that has data with multiple features of x_i to predict a target variable y_i . Most authors use XGBoost for different regression and classification problems due to the speed and prediction accuracy of the algorithm.

Extreme Gradient Boosting (XGBoost) is one of the efficient [19] algorithms in the gradient descent that has a linear model algorithm and tree learning algorithm. It is faster than other gradient descent algorithms because of the parallel computation on a single machine. This paper chooses the XGBoosting algorithm for experiments to predict the target variable daily rainfall intensity using various input or dependent environmental variables. XGBoost is a powerful algorithm that is fast learning through parallel and distributed computing and offers efficient memory usage that produces a robust solution.

Methodology

Data collection

For this study, the raw data were collected from the regional meteorological station at Bahir Dar City, Ethiopia. Ten data features such as year, month, date, evaporation, sunshine, maximum temperature, minimum temperature, humidity, wind speed, and rainfall were included. The meteorology station records the values of the environmental variable every day for each year directly from the devices in the station. Then, the data were recorded in the Microsoft Excel file tabular format. The year and the days of the month were arranged in the row of tables related to environmental variables in the column of the table.

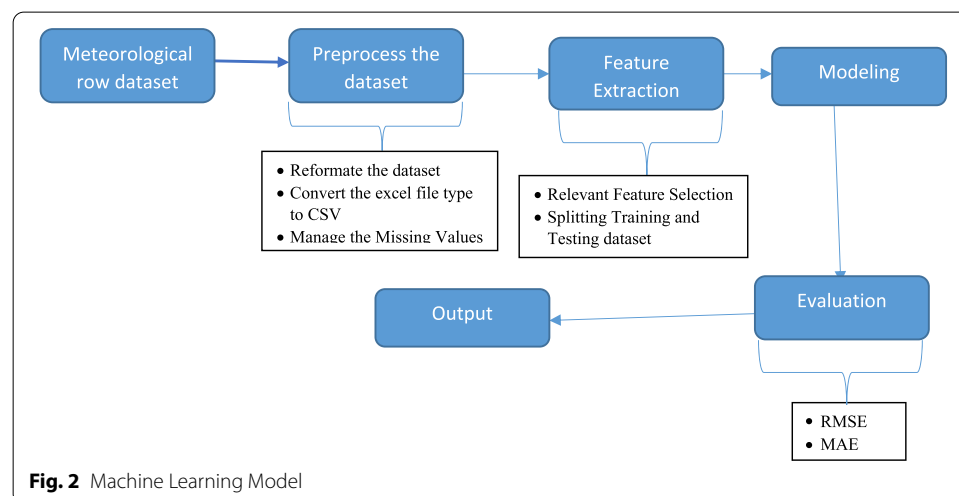
The raw data recorded at the station for 20 years (1999–2018) were used for the study.

Data preprocessing

The data preprocessing step included the data conversion, manage missing values, categorical encoding, and splitting dataset for training and testing dataset. A total of 20 years (1999–2018) data were collected from the meteorology office. Since the data were raw, they contained missing values, and wrongly encoded values so that the missing values of the target variable were removed and the other features were filled using the mean of the data.

In the meteorology office, the raw data were also arranged in a year based and the attributes in rows that need to combine and rearrange features in columns. Thus, data were converted from excel data to CSV data.

Encoding the dataset was performed and then the dataset was prepared for the experiment. The important features for rainfall prediction were selected and the dataset splitting as 80% for training and 20% for testing were considered as an input for the model.



Model

In this paper, the rainfall was predicted using a machine learning technique. Three machine learning algorithms such as Multivariate Linear Regression (MLR), Random Forest (RF), and gradient descent XGBoost were analyzed which took input variables having moderately and strongly related environmental variables with rainfall. The better machine learning algorithm was identified and reported based on the performance measure using RMSE and MAE (Fig. 2).

Measuring performance

Pearson correlation was used to measure the strength of the relationship between two variables. The two variables can be positively or negatively correlated and no relationship between the two variables if the Pearson correlation coefficient is zero. The Pearson correlation coefficient model is mathematically described as:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

where r_{xy} is the Pearson correlation coefficient, $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ are paired data consisting of n pairs and \bar{x} and \bar{y} are mean of x and y respectively.

To show the relevant features of the environmental variables to predict daily rainfall intensity, the following Pearson coefficient ranges and interpretations are used as shown in Table 1.

The machine learning algorithms take the input data features which are selected using the Pearson correlation coefficient as relevant features.

The rainfall prediction performance of each machine learning algorithm that was used in this study was measured using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) to compare which machine learning algorithms outperform better than others. RMSE and MAE were two of the most common metrics used to measure accuracy for continuous variables. The MAE measures the average magnitude of the errors in a set of forecasts and the corresponding observation, without considering their direction.

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j|$$

Table 1 Pearson coefficient ranges and interpretations

Pearson coefficient r	Interpretation
0.00 < 0.10	Negligible
0.10 < 0.20	Weak
0.20 < 0.40	Moderate
0.40 < 0.60	Relatively strong
0.60 < 0.80	Strong
0.80 < 1.00	Very strong

The RMSE is a quadratic scoring rule which measures the average magnitude of the error. It's the square root of the average of squared differences between prediction and actual observation.

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2}$$

RMSE gives a relatively high weight to large errors. This means the RMSE is most useful when large errors are particularly undesirable. The MAE and the RMSE can be used together to diagnose the variation in the errors in a set of forecasts. The RMSE will always be larger or equal to the MAE; the greater difference between them, the greater the variance in the individual errors in the sample. If the RMSE = MAE, then all the errors are of the same magnitude.

Findings

The main objective of this study was to identify the relevant atmospheric features that cause rainfall and predict the intensity of daily rainfall using machine learning techniques. Consequently, the research findings are summarized below.

To choose the environmental variables that correlate with the rainfall, the Pearson correlation was analyzed on the environmental variables presented in Table 1 above. Since the dataset is large, the variables that correlate greater than 0.20 with rainfall were considered as the participant environmental features to the experiment for rainfall prediction. Hence, to predict the amount of daily rainfall, the results of environmental attributes relevant to daily rainfall prediction like Evaporation, Relative Humidity, Sunshine, Maximum Daily Temperature, and Minimum Daily Temperature are shown in Table 2.

The Pearson Correlation coefficient experimental results on the given data showed that the attributes such as year, month, day, and wind speed had no significant impact on the prediction of rainfall. This paper took environmental values which had a correlation coefficient greater than 0.2 and analyzed the rainfall prediction. The highly correlated environmental features for rainfall prediction were relative humidity and the daily sunshine which measured the Pearson coefficient of 0.401 and 0.351 respectively.

Table 2 Environmental features and their Pearson coefficient value

Features	r
Year	0.012
Month	0.101
Day	0.017
Evaporation	0.279
Relative humidity	0.401
Max daily temperature	0.296
Min daily temperature	0.204
Sunshine	0.351
Wind speed	0.046
Daily rainfall	1.000

Table 3 Performance measurements

Algorithms	MAE	RMSE
Random forest	4.49	8.82
MLR	4.97	8.61
XGBoost	3.58	7.85

The machine learning model used the selected environmental features as an input for the algorithms. The regression models were implemented in python and the performances of the MLR, RF, and XGBoost were measured using MAE and RMSE.

In Table 3 above, the comparison of results of the three algorithms such as the MLR, RF, and XGBoost was made. The performance results indicated that XGBoost Gradient descent outperformed MLR and RF. The MAE and RMSE values of the XGBoost gradient descent algorithms were 3.58 and 7.85 respectively so that The XGBoost algorithm predicted the rainfall using relevant selected environmental features better than the RF and the MLR.

Discussion

The environmental features used in this study taken from the meteorological station collected by measuring devices are analyzed their relevance on the impact of rainfall and selected the relevant features based on experiment result of Pearson correlation values as shown in Table 2 for the daily rainfall prediction. This paper took environmental features which had a correlation coefficient greater than 0.2 and analyzed the rainfall prediction. Similarly, Manandhar et al. [7] identifies the five important environmental features such as Temperature, Relative Humidity, Dew Point, Solar Radiation, precipitable water vapor using a degree of correlation among each feature. According to the experiment result of the study, a high negative correlation coefficient of around -0.9 is observed between Temperature and Relative Humidity. The researcher Prabakaran et al. [15] used the year, temperature, cloud cover and year attribute for the experiment without analyzing the relationship between environmental features, and Gnanasankaran and Ramaraj, [14] did not show the impact of environmental features on rainfall rather used the monthly and annual rainfall data to predict the average yearly rainfall.

This study used the relevant environmental feature to train and test the three machine learning models such as RF, MLR, and XGBoost for the daily rainfall amount prediction. The performance of these machine learning models was measured using MAE and RMSE. The MAE of RF, MLR, XGBoost are 4.49, 4.97, and 3.58, and the RMSE is 8.82, 8.61, and 7.85 respectively. Similarly, the researcher Manandhar et al. [7] used data-driven machine learning algorithms to predict the annual rainfall using the selected relevant environmental features and recorded an overall accuracy of 79.6%. The researcher considered the attributes to predict the amount of yearly rainfall amount by taking the average value of temperature, cloud cover, and rainfall for a year as an input. The correlation analysis between attributes was not assessed. The average error percentage of the yearly rainfall prediction using modified linear regression was 7%. The researcher Gnanasankaran and Ramaraj [14],

did not show the impact of environmental features on rainfall. The research took the monthly and annual rainfall for the prediction of rainfall and measures the performance using RMSE which was 0.1069 and MAE which was 0.0833 using multiple linear regression.

Hence, this study assessed the impact of environmental features on the daily rainfall intensity using the Pearson correlation and selected the relevant environmental variables. The relevant features are used as an input for the daily rainfall amount prediction machine learning models and the performance of the models are measured using MAE and RMSE.

Conclusion

Rainfall Prediction is the application area of data science and machine learning to predict the state of the atmosphere. It is important to predict the rainfall intensity for effective use of water resources and crop production to reduce mortality due to flood and any disease caused by rain. This paper analyzed various machine learning algorithms for rainfall prediction. Three machine learning algorithms such as MLR, RF, and XGBoost were presented and tested using the data collected from the meteorological station at Bahir Dar City, Ethiopia.

The relevant environmental features for rainfall prediction were selected using the Pearson correlation coefficient. The selected features were used as the input variables for the machine learning model used in this paper. A comparison of results among the three algorithms (MLR, RF, and XGBoost) was made and the results showed that the XGBoost was a better-suited machine learning algorithm for daily rainfall amount prediction using selected environmental features. The accuracy of the rainfall amount prediction may increase if the sensor data is incorporated for the study. But the sensor data was not considered in this study.

The Rainfall prediction accuracy can be improved using sensor and meteorological datasets with additional different environmental features. Hence, in future work, big data analysis can be used for rainfall prediction if the sensor and meteorological datasets are used for the daily rainfall amount prediction study.

Abbreviations

XGBoost: Extreme Gradient Boosting; MLR: Multivariate Linear Regression; RF: Random Forest; RMSE: Root Mean Squared Error; MAE: Mean Absolute Error; SVM: Support Vector Machine; DT: Decision Tree.

Acknowledgements

We gratefully acknowledge the North West of Ethiopia Meteorology Agency for providing meteorological data, valuable information, and kind help for the completion of this study.

Authors' contributions

CML designed and coordinated this research, drafted the manuscript, and experiment. CML and HAM carried out the data collection and data analysis. Both the authors read and approved the final manuscript.

Funding

There are no funding organizations or individuals.

Availability of data and materials

The raw data collected from the North West of Ethiopia Meteorology Agency is available for researchers if it is requested and the materials that the authors used are available at the authors' hands.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 11 August 2021 Accepted: 23 November 2021

Published online: 07 December 2021

References

- Ehsan MA. Seasonal predictability of Ethiopian Kiremt rainfall and forecast skill of ECMWF's SEAS5 model. *Climate Dynamics*. 2021; 1–17.
- Kusiak A, Verma AP, Roz E. Modeling and prediction of rainfall using radar reflectivity data: a data-mining approach. *IEEE Trans Geosci Remote Sens*. 2013;51:2337–42.
- Chowdari KK, Girisha R, Gouda KC. A study of rainfall over India using data mining. In 2015 International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT). IEEE: New York. 2015; pp. 44–47.
- Namitha K, Jayapriya A, SanthoshKumar G. Rainfall prediction using artificial neural network on map-reduce framework. *ACM*. 2015. <https://doi.org/10.1145/2791405.2791468>.
- Tharun VP, Prakash R, Devi SR. Prediction of Rainfall Using Data Mining Techniques. In 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT). IEEE Xplore. 2018; pp. 1507–1512.
- Zainudin S, Jasim DS, Bakar AA. Comparative analysis of data mining techniques for Malaysian rainfall prediction. *Int J Adv Sci Eng Inform Technol*. 2016;6(6):1148–53.
- Manandhar S, Dev S, Lee YH, Meng YS, Winkler S. A data-driven approach for accurate rainfall prediction. *IEEE Trans Geosci Remote Sens*. 2019;5(11):9323–31.
- Arnav G, Kanchipuram Tamil Nadu. Rainfall prediction using machine learning. *Int J Innovative Sci Res Technol*. 2019. 56–58.
- Aswin S, Geetha P, Vinayakumar R. Deep learning models for the prediction of rainfall. In 2018 International Conference on Communication and Signal Processing (ICCSP). IEEE: New York. 2018; pp. 0657–0661.
- Zeelan BCMAK, Bhavana N, Bhavya P, Sowmya V. Rainfall prediction using machine learning & deep learning techniques. *Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020)*. Middlesex University: IEEE Xplore. 2020; pp. 92–97.
- Vijayan R, Mareeswari V, Mohankumar P, Gunasekaran G, Srikar K, (JUNE,). Estimating rainfall prediction using machine learning techniques on a dataset. *Int J Sci Technol Res*. 2020;9(06):440–5.
- Chaudhari MM, Choudhari DN. Study of various rainfall estimation & prediction techniques using data mining. *Am J Eng Res*. 2017;6(7):137–9.
- Thirumalai C, Harsha KS, Deepak ML, Krishna KC. Heuristic prediction of rainfall using machine learning techniques. In 2017 International Conference on Trends in Electronics and Informatics (ICEI). IEEE: New York. 2017; pp. 1114–1117.
- Gnanasankaran N, Ramaraj E. A multiple linear regression model to predict rainfall using Indian meteorological data. *Int J Adv Sci Technol*. 2020;29(8):746–58.
- Prabakaran S, Kumar PN, Tarun PSM. Rainfall prediction using modified linear regression. *ARPN J Eng Appl Sci*. 2017;12(12):3715–8.
- Balan MS, Selvan JP, Bisht HR, Gadgil YA, Khaladkar IR, Lomte VM. Rainfall prediction using deep learning on highly non-linear data. *Int J Res Eng Sci Manage*. 2019;2(3):590–2.
- Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Comput Sci*. 2021;2(6):1–20.
- Sarker IH. Machine learning: algorithms, real-world applications and research directions. *SN Comput Sci*. 2021;2(3):1–21.
- Srinivas AST, Somula R, Govinda K, Saxena A, Reddy PA. Estimating rainfall using machine learning strategies based on weather radar data. *Int J Commun Syst*. 2020;33(13):1–11.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)