International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 01 Jan 2024, pp: 506-509 www.ijaem.net ISSN: 2395-5252

# Machine Learning-Based Rainfall Forecasting

Suraj B S, C Chethana, Chetan kumar V, Sindhu B S

*Department of Computer Engineering and Engineering, PESCE,Mandya*
*Department of Computer Engineering and Engineering PESCE,Mandya*
*Department of Computer Engineering and Engineering PESCE,Mandya*
*Department of Computer Engineering and Engineering PESCE,Mandya*

Date of Submission: 20-01-2024 Date of Acceptance: 30-01-2024

**ABSTRACT**— Because of the unpredictable nature of rainfall patterns and global climatic fluctuations, rainfall prediction is a difficult undertaking. Forecasts of the amount of rain fall can avoid floods and even aid in agricultural production of crops. Rainfall prediction is made easier by machine learning, which makes use of obscure patterns found in past weather data. This paper uses a machine learning classification algorithm to predict rainfall. Of the total data, 80% is the training data and 20% is the tested data. Analysis is done on the number of actual and expected forecasts. It is discovered that the employed classification model is 80% accurate. The accuracy metrics of numerous other machine learning methods are also contrasted.

**Keywords**— classification;; machine learning; rainfall prediction;

## I. INTRODUCTION

Precipitation forecasting is an essential part of weather forecasting since erratic precipitation can lead to overflow or dry conditions, preventing floods and droughts while safeguarding infrastructure and human lives. Climate change is projected to cause floods and other extreme events to occur more frequently and with disastrous consequences [1]. Furthermore, as India is an agricultural country, forecasting precipitation is crucial to the growth of crops. It warns people to prepare ahead of time and shield their crops from rain. There are three categories for weather forecasts: short-, medium-, and long-term forecasts. Forecasts for a few hours are related to short-term forecasts. Forecasts for the medium term span four to ten days, while those for the long term span more than ten days. When it comes to flood forecasting, short- and medium-term projections are useful[2].

Because rainfall is erratic, predicting it is seen to be a difficult undertaking. One of the things that contributes to this unpredictability, threatens communities, and keeps them growing is climate change[3]. Therefore, it is essential to forecast rainfall accurately. When compared to conventional models that rely on statistics, it is discovered that intelligent prediction models produce better results. However, based on accuracy, these intelligent models' output can still be enhanced. This research uses a machine learning approach to predict rainfall. Given that numerous techniques exist for machine learning the paper focus on the classification method that, based on the model's training, distinguishes all of the provided data and contributes to a better forecast. In this paper, other supervised-learning-based machine learning approaches are also compared for better accuracy. It can be concluded that the used classification method outperformed all other classifiers, demonstrating the method's successful implementation for precipitation prediction.

## II. RELATED WORK

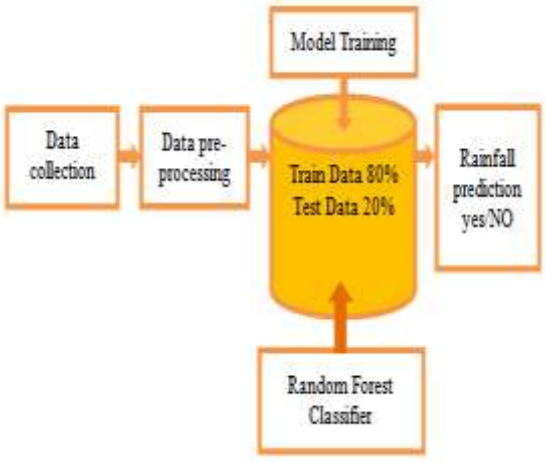
Multiple linear regression and neural networks are used to predict rainfall, as Patil [7] demonstrates. Maqaleh[8] discusses two approaches for forecasting: one is based on statistical models, and the other is developed using an artificial neural network. Box-Jenkins model is utilized as first technique which is utilized to compute time series. The second approach, which improves the weights of a back propagation neural network by utilizing adaptive slope and momentum parameter, is a modified artificial neural network model. Rahman[9] utilizes AI strategies, for example, Choice Tree, Gullible Bayes, K-Closest Neighbors, and Backing Vector Machines their expectation precision are reinforced by fluffy rationale. Machine learning and artificial neural networks are compared by Dutta[10]. LASSO regression is used in machine learning, which is more accurate than the artificial neural network method... Gupta[13] broke down that brain networks convey further developed impacts than any of different calculations with 10 neurons in secret layer. K. R. uses clustering and classification data mining

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 01 Jan 2024, pp: 506-509 www.ijaem.net ISSN: 2395-5252

methods to predict rainfall. Mishra[15] executes Feed Forward Brain Organization (FFNN) for anticipating the precipitation. Nigam [16] showed that Straightforward Repetitive Brain Organization accomplishes further developed outcome on precipitation expectation in spite of the fact that LSTM is better for temperature expectation. According to Kadam [5], using machine learning techniques, it is likely to accurately predict rainfall.. system and shuffled frog leaping algorithm. The model was implemented on data taken from two different rivers and it was concluded that the model gave better results than simple adaptive neuro –fuzzy inference system.

## III. PROPOSED WORK

According to Gupta's analysis [13], neural networks outperform all other algorithms with ten neurons in the hidden layer in terms of impact. K. R.[14] uses data mining approaches for rainfall prediction, including clustering and classification. Feed Forward Neural Network (FFNN) is used by Mishra[15] to forecast rainfall. Nigam [16] shown that while LSTM is superior for temperature prediction, Simple Recurrent Neural Networks perform better when it comes to rainfall prediction. Kadam [5] suggests that using machine learning techniques would probably result in an efficient Rainfall prediction. training is done and the test set is one which helps in making predictions. The classification is done using Random Forest classifier which gives outcome in form of YES and NO for the forecast of rain and this forecast is compared with the actual rain conditions in the study area. The proposed approach gives an overall 80% model accuracy.




Rainfall Prediction Model

*A.* Random Forest Classifier

The Random Forest classifier constructs its building pieces from decision trees. From the provided dataset, a random sample is selected, and a decision tree is built for each sample. The decision tree's predictions for each sample are based on the least amount of correlation between them. Larger datasets yield higher accuracy results from the random forest classifier. In addition, the Random Forest Classifier requires less training time than previous methods. The Python code that was used to construct the Random Forest Classifier is seen below. When the suggested model's Random Forest classifier is being implemented, the value of Since more trees result in better accuracy and the random state is set to 0, the number of estimators (n) is chosen to be 140.

## IV. RESULTS AND DISCUSSIONS

Finding the differences between the actual forecast and the projected forecast is made easier by using the Rainfall prediction model, which is implemented shows the result that the model produced. Four scenarios are taken into consideration while counting the actual and predicted rainfall. When there is absolutely no rainfall seen, Case 1 holds that both the actual and predicted quantities are erroneous. Case 2 presents the actual rainfall as erroneous and the anticipated rainfall as occurring. It illustrates how, despite the weather forecast's indications of rain, there was really no rainfall seen. In Case 3, the forecast is shown as false and the actual forecast as true. This indicates that rain fell on the days when no rain estimate was made. Case 4 accepts as true both the actual and predicted results. In this case, the observation and the actual rainfall-causing meteorological conditions match exactly.

The likelihoods for the actual and predicted forecasts for each of the four scenarios are detailed in Table 1. It is evident that 70% of predictions and actuals agree that there won't be any rain tomorrow. In the event that there is no rain, even if the rain forecast was verified, there are just 4% of chances. Eleven percent of the chances happened while it was really raining but it wasn't expected to rain tomorrow. Nonetheless, 15% of the time, the next day was predicted to be a wet day and the exact amount of rain fell as expected. Thus, 80% of the total possibilities showed accurate findings when comparing the predicted and actual amounts of rain for tomorrow. As a consequence, the model that was used produced the best results when all significant factors that affect rainfall and assistance were taken into account.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 01 Jan 2024, pp: 506-509 www.ijaem.net ISSN: 2395-5252

| Chances | Actual Rain Tomorrow | Predicted Rain Tomorrow |
|---|---|---|
| 70% | NO | NO |
| 4% | NO | YES |
| 11% | YES | NO |
| 15% | YES | YES |

A. Comparison of results

In order to determine whether results will improve, a machine learning approach compares several supervised learning classifiers, such as Naive Bayes, support vector machines, K nearest neighbors, and Random Forest, based on accuracy. Based, it can be concluded that the Random Forest Classifier has the highest accuracy of 80%, followed by K nearest neighbor (KNN) with 75% accuracy and Naive Bayes with 76% accuracy. The support vector machine (SVM) has the lowest accuracy, at 43%.

## V. CONCLUSION

Forecasting rainfall is crucial for organizing food production, preventing floods, and efficiently managing water resources. This article uses a machine learning classification approach to improve rainfall prediction forecasts. The Random Forest classifier was used here because, in contrast to other methods, it can tolerate missing values and a big dataset with less training time because the majority of rainfall datasets available had irregular patterns. With 80% accuracy, it has produced the best outcome. When compared to alternative approaches, the suggested framework produced the best results, demonstrating an improvement in the findings. Future research will use this method to justify its use with a wider range of datasets.

## REFERENCES

[1]. C. L. Wu and K. W. Chau, "Prediction of rainfall time series using modular soft computingmethods," Eng. Appl. Artif. Intell., vol. 26, no. 3, pp. 997–1007, 2013, doi: 10.1016/j.engappai.2012.05.023.

[2]. S. Aftab, M. Ahmad, N. Hameed, M. S. Bashir, I. Ali, and Z. Nawaz, "Rainfall prediction using data mining techniques: A systematic literature review," Int. J. Adv. Comput. Sci. Appl., vol. 9, no. 5, pp. 143–150, 2018, doi: 10.14569/IJACSA.2018.090518.

[3]. A. Y. Barrera-animas et al., "Machine Learning with Applications Rainfall prediction : A comparative analysis of modern machine learning algorithms for time-series forecasting," Mach. Learn. with Appl., vol. 7, no. October 2021, p. 100204, 2022, doi: 10.1016/j.mlwa.2021.100204.

[4]. M. A. Nayak and S. Ghosh, "Prediction of extreme rainfall event using weather pattern recognition and support vector machine classifier," Theor. Appl. Climatol., vol. 114, no. 3–4, pp. 583–603, 2013, doi: 10.1007/s00704-013-0867-3.

[5]. D. Salgaonkar and B. Chaudhari, "Survey on Rainfall Prediction System," no. 09, pp. 9–11, 2021.

[6]. M. Hasan et al., "ScienceDirect Search and rescue operation in flooded areas : A survey on emerging sensor networking-enabled IoT-oriented technologies and applications," Cogn. Syst. Res., vol. 67, pp. 104–123, 2021, doi: 10.1016/j.cogsys.2020.12.008.

[7]. Shakib Badarpura, Abhishek Jain, Aniket Gupta, and Deepali Patil, "Rainfall Prediction using Linear approach Neural Networks and Crop Recommendation based on Decision Tree," Int. J. Eng. Res., vol. V9, no. 04, pp. 394–399, 2020, doi: 10.17577/ijertv9is040314.

[8]. B. M. Al-Maqaleh, A. A. Al-Mansoub, and F. N. Al-Badani, "Forecasting using Artificial Neural Network and Statistics Models," Int. J. Educ. Manag. Eng., vol. 6, no. 3, pp. 20–32, 2016, doi: 10.5815/ijeme.2016.03.03.

[9]. A. U. Rahman et al., "Rainfall Prediction System Using Machine Learning Fusion for Smart Cities," Sensors 2022, Vol. 22, Page 3504, vol. 22, no. 9, p. 3504, May 2022, doi: 10.3390/S22093504.

[10]. K. Dutta and G. P*, "Rainfall Prediction using Machine Learning and Neural Network," Int. J. Recent Technol. Eng., vol. 9, no. 1, pp. 1954–1961, 2020, doi: 10.35940/ijrte.a2747.059120.

[11]. C. Z. Basha, N. Bhavana, P. Bhavya, and V. Sowmya, "Rainfall Prediction using Machine Learning Deep Learning Techniques," Proc. Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2020, no. Icesc, pp. 92–97, 2020, doi: 10.1109/ICESC48915.2020.9155896.

[12]. K. R. B and G. A, "Rainfall Prediction Using Data Mining Techniques - A Survey," vol. 83, no. 8, pp. 23–30, 2013, doi: 10.5121/csit.2013.3903.

[13]. D. Gupta and U. Ghose, "A comparative study of classification algorithms for forecasting rainfall," 2015 4th Int. Conf. Reliab. Infocom Technol. Optim. Trends

Futur. Dir. ICRITO 2015, pp. 1–6, 2015, doi: 10.1109/ICRITO.2015.7359273.

[14]. K. R. B and G. A, "Rainfall Prediction Using Data Mining Techniques - A Survey," vol. 83, no. 8, pp. 23–30, 2013, doi: 10.5121/csit.2013.3903.

[15]. N. Mishra, H. K. Soni, S. Sharma, and A. K. Upadhyay, "A comprehensive survey of data mining techniques on time series data for rainfall prediction," J. ICT Res. Appl., vol. 11, no. 2, pp. 167–183, 2017, doi: 10.5614/itbj.ict.res.appl.2017.11.2.4.

[16]. J. Ranjani, V. K. G. Kalaiselvi, A. Sheela, D. Deepika Sree, and G. Janaki, "Crop Yield Prediction Using Machine Learning Algorithm," Proc. 2021 4th Int. Conf. Comput. Commun. Technol. ICCCT 2021, pp. 611–616, 2021, doi: 10.1109/ICCCT53315.2021.9711853.

[17]. W. M. Ridwan, M. Sapitang, A. Aziz, K. Faizal, A. Najah, and A. El-shafie, "Rainfall forecasting model using machine learning methods : Case study," Ain Shams Eng. J., vol. 12, no. 2, pp. 1651–1663, 2021, doi: 10.1016/j.asej.2020.09.011.

[18]. E. A. Hussein, M. Ghaziasgar, C. Thron, and M. Vaccari, Rainfall Prediction Using Machine Learning Models : Literature Survey Rainfall Prediction Using Machine Learning Models : Literature Survey, no. January. 2022.

[19]. B. Mohammadi and S. Mehdizadeh, "Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm," Agric. Water Manag., no. September 2019, p. 106145, 2020, doi: 10.1016/j.agwat.2020.106145.

[20]. B. Mohammadi et al., "Adaptive neuro-fuzzy inference system coupled with shuffled frog leaping algorithm for predicting river streamflow time series," Hydrol. Sci. J., vol. 0, no. 0, p. 1, 2020, doi: 10.1080/02626667.2020.1758703.

[21]. T. O. Muslim et al., "Investigating the Influence of Meteorological Parameters on the Accuracy of Sea-Level Prediction Models in Sabah , Malaysia," pp. 1–18, 2023.

[22]. "Rain in Australia | Kaggle." https://www.kaggle.com/datasets/jsphyg/weather-dataset-rattle-package?resource=download (accessed Sep. 06, 2022).