A Review Paper on Machine Learning Based Models for Rainfall Pattern Prediction

[1]Nilesh Dhannaseth, [2]Dr. Sanjay Yedey
[1]Research Scholar, [2]Associate Professor
[1,2]P.G. Department of Computer Science & Technology
[1,2]DCPE, HVPM, Amravati, India

Abstract-Rainfall is a pivotal rainfall parameter in the environment of India. vaticination of downfall can effectively prop the decision-making process for husbandry and natural disaster operation of the country. still the chaotic nature of downfall due to climate change has made the task of downfall vaticination challenging through traditional statistical models. In this study, we dissect the performance of machine literacy algorithms Decision Tree (DT), K- Nearest Neighbours (KNN), Random Forest (RF), Extreme Gradient Boosting (XGB), Light grade Boosting (LGB) and Multi-Layered Perceptron (MLP) in prognosticating diurnal downfall as both retrogression and bracket. The models were trained with and without point selection and/ or slice ways (for bracket). During training10-fold cross confirmation and hyper parameter tuning was performed on the train set and latterly the named models were applied to the test set for evaluation.

Index Terms—Rainfall Pattern Prediction, Machine Learning, Decision Tree, Classification, Regression

I. INTRODUCTION

A better soothsaying model is demanded for an early warning that can limit threats to life and property while also better managing agrarian fields because inordinate and irregular rainfall can have a variety of effects, including crop loss and property damage. Farmers stand to gain the most from this projection, and water resources may be utilized more effectively. Forecasting rainfall is a challenging task, and the results must be accurate. Based on meteorological information including temperature, humidity, and pressure, for the purpose of forecasting rainfall, various physical tools are available. Because traditional methods are less effective, machine learning algorithms like Decision Trees, Multivariate Linear Regression, and Random

Forest are used to produce accurate results. By looking at historical rainfall data and predicting rainfall for upcoming seasons, it may be done with ease.
Downfall is a vital rainfall miracle for any region of the world, especially for a country like India where utmost of the population is still directly or laterally dependent on husbandry. In the environment of India, downfall not only plays a pivotal part in the country's crop yield but also is associated with natural disasters similar as cyclones, storms and cataracts. A dependable vaticination model for downfall would clearly aid the policy makers of the country in terms of husbandry and natural disaster operation. Due to rapid-fire climate change the rainfall parameters of India including downfall has come relatively changeable (1). therefore, accurate vaticination of downfall has come a challenge for the traditional statistical styles. In this paper we estimate different machine learning algorithms Decision Tree (DT), K- Nearest Neighbours (KNN), Random timber (RF), Extreme Gradient Boosting (XGB), Light Gradient Boosting (LGB) and Multi-Layered Perceptron (MLP) in prognosticating downfall of India for both retrogression and bracket. Improving the accuracy of machine learning techniques on weather forecasting has been the primary concern of many researchers over the last two decades. Some of the related studies are discussed here.

In [9], researchers presented an ANN-based technique to predict atmospheric conditions. The dataset used for prediction consisted of various weather attributes, e.g., humidity, temperature, and wind speed.

In [10], researchers proposed a hybrid method for rainfall forecasting by integrating feature extraction and prediction techniques. The dataset used for the experiment was obtained from the National Oceanic and Atmospheric Administration (NOAA); it spanned more than 50 years and consisted of various weather features such as humidity, pressure, temperature, and wind speed. A Neural Network was used to classify the instances into low, medium, and highclasses based on a pre-defined training set.

In [11], researchers presented a data-intensive model for rainfall prediction using a Bayesian modeling approach. For the experiment, the dataset was collected from the Indian Meteorological Department, and from 36 attributes, the 7 most relevant attributes were selected. Before the prediction, pre-processing and transformation steps were performed for smooth processing. The proposed approach showed good accuracy for rainfall prediction, using moderate computing resources compared to meteorological centers using high-performance computing power for weather predictions.
Using emitted long wave radiations, average world temperatures, and sunspots from Tamil Nadu, Samuel and Raajalakshmi [2] employed multiple linear regression to forecast monsoon rainfall. They gathered information from Chennai's Indian Meteorological Department for 110 years.

A forecasting model was created by Paras and Sanjay [4] using mathematical regression. The algorithm uses three years' worth of weather data to forecast maximum and minimum temperatures 15 to 45 weeks in the future.

II. RELATED WORKS

Various methods for rainfall analysis are examined. These include statistical techniques such as frequency analysis, time series analysis, and spatial interpolation. Moreover, the review explores the utilization of advanced technologies like remote sensing, machine learning, and data mining for robust and accurate analysis.

Climate change's impact on rainfall patterns is a significant focus. The review investigates how shifting climate conditions alter rainfall distribution, intensity, and seasonality. This section underscores the importance of adapting analysis techniques to account for these changes.

Applications of rainfall analysis are explored, spanning flood prediction, water resource management, urban planning, and agricultural strategies. Real-world case studies are used to illustrate the practical implications of rainfall analysis in diverse scenarios.

Challenges are acknowledged, such as data quality issues, uncertainties, and the need for improved spatial and temporal resolution. The review also points toward emerging technologies and methodologies that hold promise for enhancing the accuracy and scope of rainfall analysis.

In conclusion, the review underscores the multifaceted nature of rainfall analysis, demonstrating its value in understanding natural variability, predicting extreme events, and informing decision-making across various sectors. It reinforces the need for continued research and innovation in this field to address the evolving challenges posed by a changing climate.

Over the past two decades, many researchers have focused on increasing the precision of machine learning techniques used in weather forecasting. Here, some related studies are covered. Researchers described an ANN-based method to forecast atmospheric conditions in [9]. A variety of meteorological characteristics, such as humidity, temperature, and wind speed, were included in the dataset utilized for the prediction.

Researchers in [10] suggested a hybrid strategy for predicting rainfall by combining feature extraction and prediction algorithms. The National Oceanic and Atmospheric Administration (NOAA) provided the dataset for the experiment, which included data on humidity, pressure, temperature, and wind speed over a period of more than 50 years. Based on a predetermined training set, a Neural Network was used to categories the cases into low, middle, and high classes.

In [11], researchers presented a data-intensive model for rainfall prediction using a Bayesian modeling approach. For the experiment, the dataset was collected from the Indian Meteorological Department, and from 36 attributes, the 7 most relevant attributes were selected. Before the prediction, pre-processing and transformation steps were performed for smooth

processing. The proposed approach showed good accuracy for rainfall prediction, using moderate computing resources compared to meteorological centers using high-performance computing power for weather predictions.

Research papers	Methods	Advantages	Gap
AyishaSiddiqua L, Senthilkumar N C. "Heavy Rainfall Prediction using Gini Index in Decision Tree".	Decision Tree, Gini Index	decreases the complexity in computation	Do not take care of other wheather parameters
Wanie M.Ridwan. "Rainfall forecasting model using machine learning methods: Case study Terengganu, Malaysia	Bayesian Linear Regression (BLR), Boosted Decision Tree Regression	Provide acceptable result	Exclusion of different climate change scenarios.
Cmak Zeelan Basha. "Rainfall Prediction using Machine Learning & Deep Learning Techniques".	Auto-regressive integrated moving averages,Support vector machines	Here non linear data is taken into consideration	Excluded scenarios of light and wheather.

Table 1: Critical Review of Literature

III. PROPOSED METHOD

Multivariate Linear Regression:

Regression is a statistical empirical data mining technique with a broad range of applications, including business, biology, and climate prediction. It is evolving into a useful tool for society.

A supervised machine learning approach called multivariate linear regression analyses several data variables. One dependent variable and many independent variables make up a multivariate regression, which is an extension of multiple regressions. We attempt to forecast the result based on the number of independent variables.

The multivariate regression model's generalised equation is shown below: $y = \alpha 0 + \alpha 1.x1 + \alpha 2.x2 + ... + \alpha n.xn$

In this equation, n stands for the number of independent variables, $\alpha 0$-αn for coefficients, and x1=xn for the independent variable.

We can comprehend and compare coefficients across the output thanks to the multivariate model. In this case, multivariate linear regression is a superior model due to the small cost function.

As per the finding from research gaps in previous papers, following parameters can be used precipitation, sunshine, temperature, humidity, wind speed with Multivariate Linear Regression model to predict rainfall pattern.

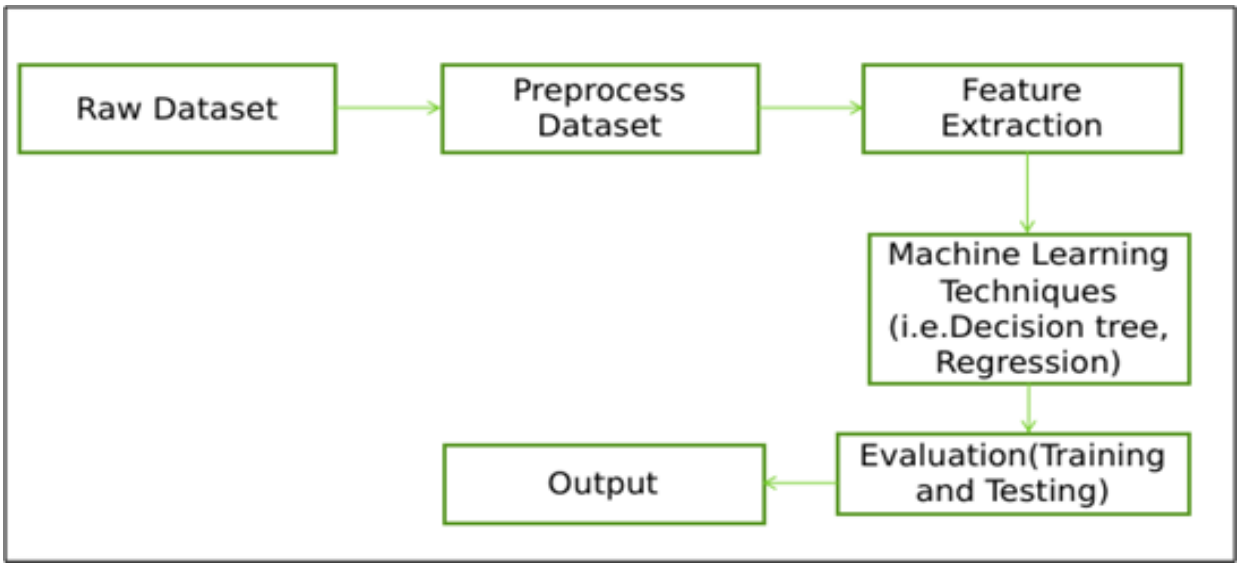

Fig.1. Architectural Flow

- Statistical Methods for Rainfall

Analysis: Review statistical techniques used to analyze rainfall data, including frequency analysis, probability distributions, and extreme value analysis. Explore approaches to identifying trends and changes in rainfall patterns over time. Discuss methods for detecting anomalies, such as droughts or extreme precipitation events.

- Climate Change and Rainfall Patterns:

Examine research on the potential impacts of climate change on rainfall patterns and variability. Discuss studies that investigate shifts in rainfall distribution, changes in seasonality, and altered frequency of extreme events due to climate change.

- Spatial and Temporal Analysis:

Explore studies that focus on spatial variations in rainfall patterns across different regions and landscapes. Discuss how temporal analysis can reveal seasonal trends, long-term variations, and cyclic patterns in rainfall data.

- Rainfall-Runoff Modeling:

Review literature related to rainfall-runoff modeling, which involves analyzing how rainfall translates into stream flow and runoff in watersheds. Discuss different modeling approaches, including conceptual models, physically-based models, and data-driven techniques.

- Applications of Rainfall Analysis:

Explore various applications of rainfall analysis, such as flood prediction, water resource management, agricultural planning, and urban infrastructure design. Provide examples of real-world projects or case studies that have utilized rainfall analysis to address specific challenges.

3.1 KEY POINTS FOR ANALYSIS

This section involves examining existing research, studies, and theories related to the analysis of rainfall patterns, trends, variability, and their implications.

- *Rainfall Characteristics and Variability:*

Discuss the fundamental characteristics of rainfall, including intensity, duration, frequency, and distribution. Review studies that have analyzed the natural variability of rainfall over different temporal and spatial scales. Explore factors that influence rainfall variability, such as geographical location, topography, and climate systems

- Rainfall Measurement and Data Sources:

Discuss different methods of rainfall measurement, including ground-based rain gauges, weather radar, and satellite-based measurements. Evaluate the strengths and limitations of each measurement method. Highlight the importance of accurate and reliable rainfall data for meaningful analysis.

- Data Challenges and Future Directions:

Address potential data limitations, such as gaps in historical records or inconsistencies in measurement methods. Discuss emerging technologies and methodologies that could enhance the accuracy and scope of rainfall analysis.

IV. CONCLUSION

Forecasting rainfall is crucial because excessive and irregular rainfall can have a variety of repercussions, such as crop loss and property damage, therefore a better forecasting model is required for an early warning that can limit risks to life and property while also better managing agricultural fields. This forecast primarily benefits farmers, and water resources may be used more efficiently.

V. FUTURE WORK

Future work in the context of the proposed evolutionary algorithm-based approach for efficient rainfall prediction.

REFERENCES

[1] Rahman, A. U., Abbas, S., Gollapalli, M., Ahmed, R., Aftab, S., Ahmad, M., Khan, M. A., & Mosavi, A. (2022, May 4). Rainfall Prediction System Using Machine Learning Fusion for Smart Cities. *Sensors*, *22*(9), 3504.

[2] R. Samuel Selvaraj and Raajalakshmi, "Statistical Method of Predicting the Northeast Rainfall of TamilNadu", Universal Journal of Environmental Research and Technology. Volume 1, Issue 4: 557-559.

[3] Kumar, V., Yadav, V. K., & Dubey, E. S. (2022, May 31). Rainfall Prediction using Machine Learning. *International Journal for Research in Applied Science and Engineering Technology*, *10*(5), 2494–2497.

[4] Paras and Sanjay Mathur "A Simple Weather Forecasting Model Using Mathematical Regression" Department of Electronics & Communication Engineering, College of Technology, G.B. Pant University of Agriculture & Technology, Pantnagar, (India) 263 145.

[5] Cmak Zeelan; Nagulla B, (2020). Rainfall Prediction using Machine Learning & Deep Learning Techniques. IEEE Xplore.

[6] Siddiqua L*, A., & N C, S. (2019, November 30). Heavy Rainfall Prediction using Gini Index in Decision Tree. *International Journal of Recent Technology and Engineering (IJRTE)*, *8*(4), 4558–4562.

[7] Ridwan, W. M., Sapitang, M., Aziz, A., Kushiar, K. F., Ahmed, A. N., & El-Shafie, A. (2021, June). Rainfall forecasting model using machine learning methods: Case study Terengganu, Malaysia. *Ain Shams Engineering Journal*, *12*(2), 1651–1663.

[8] Chowdari K., Girisha R., Gouda K. (2015). A study of Rainfall over India Using Data Mining. International Conference on Emerging Research in Electronics, Computer Science and Technology.

[9] Sawale, G.J.; Gupta, S.R. Use of Artificial Neural Network in Data Mining For Weather Forecasting. Int. J. Comput. Sci. Appl. 2013, 6, 383–387.

[10] Joseph, J. Rainfall Prediction using Data Mining Techniques. Int. J. Comput. Appl.2013, 83, 11–15.

[11] Nikam, V.B.; Meshram, B.B. Modeling rainfall prediction using data mining method: A bayesian approach. In Proceedings of the International Conference on Computational Intelligence, Modelling and Simulation, Bangkok, Thailand, 24–25 September 2013; pp. 132–136.