

Real-time rainfall challenges forecasting and prediction with machine learning.

Channabasava¹, Prathibha K N², Annapa Reddy Haarika³, Rajeswari M⁴, Rekha V⁵, Prathibha K N⁶, Reena Dudda Krishnamurty⁷, VijayKumar⁷, Rudresha S⁸

¹Department of MCA , Nitte Meenakshi institute of technology, Nitte Deemed-to-be University, Bangalore

²Department of Computer Science & Engineering, Jyothy Institute of Technology , Thathaguni, kanakapura road, Bengaluru 560082, Karnataka

³Department of Computer Science and Engineering, BMS Institute of Technology and Management, Doddaballapur Main Road, Avalahalli, Yelahanka, Bengaluru – 560119, Karnataka, India.

⁴Department of Computer Science & Engineering, Jyothy Institute of Technology , Thathaguni, kanakapura road, Bengaluru 560082, Karnataka

⁵Department of Computer Science & Engineering, Jyothy Institute of Technology , Thathaguni, kanakapura road, Bengaluru 560082, Karnataka

⁶Department of Computer Science , Atria Institute of Technology, Bengaluru, Karnataka.

⁷Department of Computer Science & Engineering, Atria institute of technology, Bengaluru, Karnataka.

⁸Department of Mechanical Engineering, University of Visvesvaraya College of Engineering K.R Circle, Bengaluru, Karnataka

Abstract

India is fundamentally an agricultural country, crop productivity and rainfall play a major role in its economy. In order to assess the crop yield, every farmer must be able to predict rainfall. Application of scientific methods and modern technologies to anticipate atmospheric patterns is known as rainfall prediction. Estimating rainfall precisely is important factor for agriculture production, water resources optimal use and development of water structure. It can predict rainfall by using various methods of data mining. Data mining methods are used to estimate rainfall mathematically. The main aim of this research is data mining techniques that are widely used for rainfall prediction. This research examines various algorithms, such as fuzzy logic, neural networks, decision trees, the K-Nearest Neighbor method, and Naive Bayes. It can assess which method forecasts rainfall with more accurately based on that comparison.

Kay words: Rainfall Prediction, Machine learning, Remote sensing Hybrid models.

1. Introductions

One of the most challenging tasks is forecasting the rainfall. Despite various techniques are proposed, forecasting rainfall precisely is still a significant challenge. India. The prospects of crop yields and water shortages are analysed in any particular year are closely monitored. The agricultural sector may experience significant variation in the seasonal rainfall. There are no known techniques available to forecast whether there will be rainfall today. factors. Even forecast from meteorological department will also be in accurate. In this article, we will explore how to create a machine-learning model which will determine probability of rainfall based on some atmospheric factors. This issue is pertaining to prediction of rainfall using machine learning as such models generally achieve good results on tasks that needed highly skilled professionals.

2. Importing Libraries and Dataset in Rain Fall Prediction.

1. Python is libraries make it easy for us to handle the data and perform typical and complex tasks with a single line of code.
2. Pandas is This library helps to load the data frame in a 2D array format and has multiple functions to perform analysis tasks in one go.
3. Numpy is Numpy arrays are very fast and can perform large computations in a very short time.
4. Matplotlib/Seaborn is This library is used to draw visualizations.
5. Sklearn is module contains multiple libraries are having pre-implemented functions to perform tasks from data preprocessing to model development and evaluation.
6. Xgboost is contains the eXtreme Gradient Boosting machine learning algorithm which is one of the algorithms which helps us to achieve high accuracy on predictions.
7. Imblearn is module contains a function that can be used for handling problems related to data imbalance.

3. Literature review

By using instrument-input-features, the machine-learning model can predict precipitation. The assessment of regression techniques in machine learning is based on how predicted values align with the actual observed values. The RMSE, R2, and MAE statistical measures evaluate the accuracy of a prediction or forecasting model. Irrespective of climate conditions or time periods, Machine learning is highly effective at forecasting rainfall. The LSTM models

showcase their excellence. Models such as Remote sensing and hybrid predictive should be used for further exploration as they are currently limited [1]. This study examined global climate factors such as Southwest Monsoon (SWM) and Northeast Monsoon (NEM) in Coimbatore region as this region experiences high rainfall during NEM. Techniques like normalisation and scaling addressed missing values keeping 70–85% of the initial data for the training set. Findings have shown that MPNN surpassed RFR's performance—achieving an accuracy of 85.55% for SWM and NEM, whereas RFR surpassed MPNN, achieving an accuracy of 86.50%. For daily rainfall from 2020 to 2023, the coefficient of determination (R^2) between predicted and observed values was 0.8 [2]. Artificial Neural Network, Logistic Regression, Support Vector Machine and Self Organizing Map. For forecasting seasonal rainfall, two frequently used models are Linear and Non-Linear models. ARIMA-Model is a Linear model while rainfall prediction through Artificial Neural Networks (ANN) can be achieved with Back Propagation NN, Cascade NN, or Layer Recurrent Networks. Artificial NN is similar to Biological Neural Networks [3]. Both a stacked auto encoder (StAE) and a sparse autoencoder (SpAE) feature an input layer for raw data, a hidden layer for extracting features, and an output layer for classification and prediction. [6]. The cities of Oda and Gotsu in Shimane Prefecture, located in south-western Japan, were used for susceptibility assessment and prediction of landslides caused by extreme rainfall as a study case [6]. This study proposes an innovative real-time rainfall prediction for smart cities using a machine learning fusion technique. The suggested framework uses four commonly used supervised machine learning techniques, i.e., decision tree, Naïve Bayes, K-nearest neighbors, and support vector machines. For more accurate rainfall predictions, the fuzzy logic method is integrated in the framework for enhancing the predictive capabilities of machine learning techniques, a process known as fusion. For forecasting, historical weather data spanning 12 years (from 2005 to 2017) for Lahore has been used. Pre-processing tasks such as cleaning and normalization were carried out on the dataset prior to the classification process. The findings indicate machine learning fusion is more efficient than other models [7]. The findings shows rain water harvesting systems with larger storage capacity situated in more humid cities can yield enhanced water-saving efficiency and reliability. Rainfall time duration plays an important role in the modelling results of RHS. The RTSL for the 12 cities vary from 6 to 21 years. The RTSL values for 12 cities show no significant correlated with mean annual rainfall ($R^2 = 0.32$, $n = 12$, $p > 0.05$) and seasonality index ($R^2 = 0.28$, $n = 12$, $p > 0.05$), but significantly correlated with variation coefficients of annual rainfall ($R^2 = 0.76$, $n = 12$, $p < 0.05$) [8].

Forecasting rainfall is essential for managing water resources including the operation of reservoirs and dams, as well as planning for water supply and distribution. By predicting rainfall accurately water supply decision-makers can make better-informed choices about water distribution, management, and conservation efforts. Precise rainfall predictions are important for mitigating the impact of disasters such as floods, landslides, and droughts. Early warning systems based on precipitation forecasts can enable communities to effectively prepare and respond, thereby minimizing the loss of life and damage to property [9,10]. Precise forecasts of rainfall can assist transportation planners in making well-informed decisions about the construction, upkeep, and safety of roads and bridges. It can also help prevent delays and accidents caused by weather. Hydroelectric power generation heavily depends on accurate precipitation forecasts that are essential for the scheduling and supervision of energy output. Rainfall also have an impact on the planning and maintenance of other forms of energy production, like solar and wind energy as it can impact their effectiveness. Forecasting rainfall is important for sustainable development, disaster risk reduction, and community welfare. It is essential for multiple industries and precise and timely predictions can help to minimise economic losses, safeguard environment and enhance quality of life [11,12]. However, inaccurate forecasting of rainfall can adversely affect many aspects of civilization including agriculture, disaster response, and the development of infrastructure. Those who depend on weather predictions should be made aware of any possible information gaps, and efforts are being made to improve forecast reliability. This review paper examines both previous and current studies for rainfall prediction through machine learning and remote sensing techniques were reviewed and analysed with the aim of investigating the novelty benefits of these predictive models.

3.Methodology

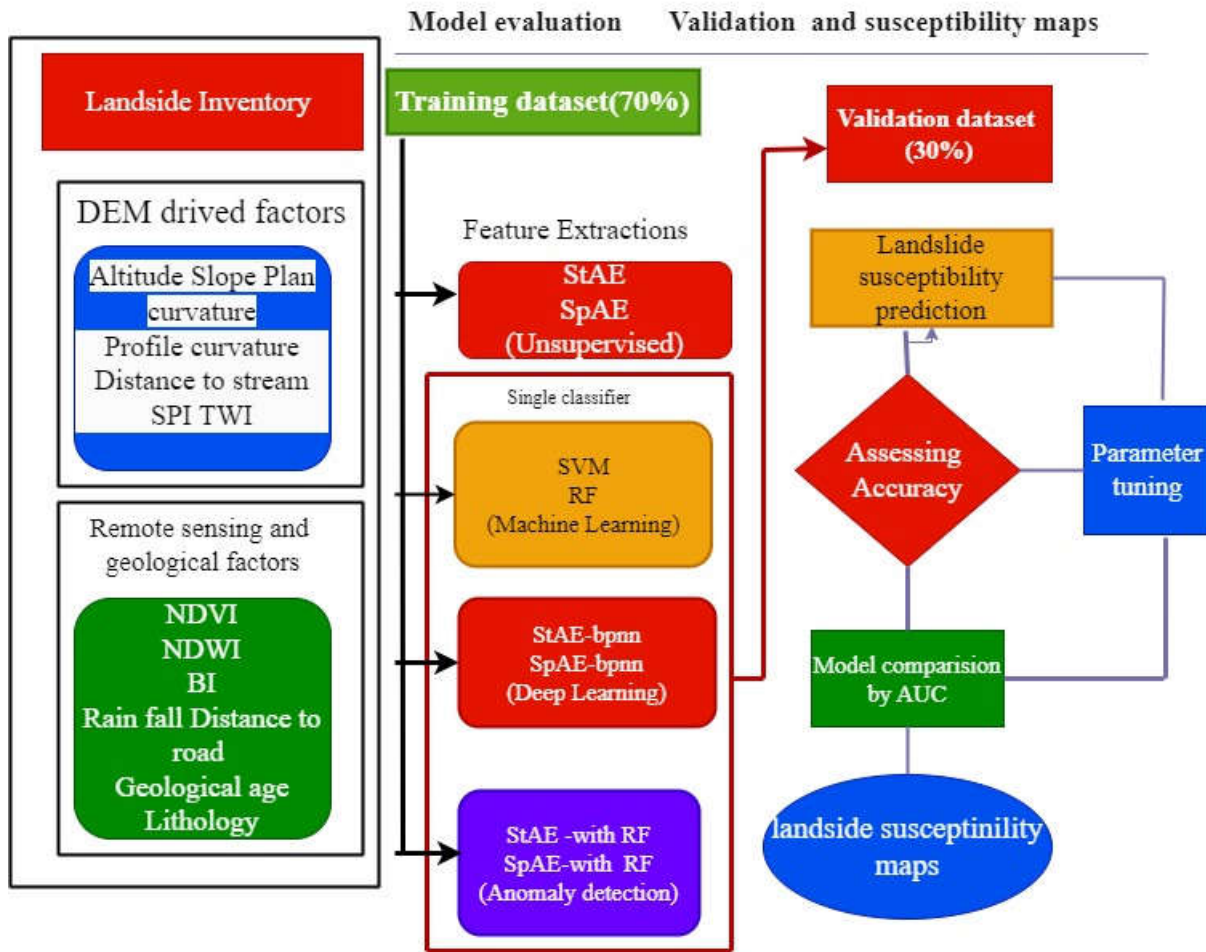


Fig:1. Methodology of rain fall prediction Machine Learning

This research was performed following these primary steps (Fig. 6): correlation analysis between the landslide inventory and factors that influence landslides using frequency ratio, (2) Prediction of landslide susceptibility through machine learning with SVM and RF models. (3) Prediction of landslide susceptibility through deep learning using StAE and SpAE with back propagation neural networks. (4) Judging the capability of StAE and SpAE when combined with machine learning based on improved prediction ratios from SVM and RF and (5) Substantiating and comparing the predictive performance using the area under the curves and the landslide susceptibility maps generated by the six models.

The samples from the landslide were created following collection and preparation of the landslide inventory map, DEM derived factors, and the geological and remote sensing factors. The samples of the landslide inventory were counted and used randomly to generate non-landslide samples. The final dataset merged the landslide and non-landslides samples with a defined label (1 and 0, respectively) for each sample. A spatial database was used to prepare fourteen landslide-affecting factors. The factors' values at each sampling location were applied,

and the resulting information was generated using R Studio. The dependent variable was converted with one-hot encoding. training (70%) and validation (30%).

The StAE and SpAE models were trained in an unsupervised way to extract features, and a set of new features was generated. These new features were used to train StAE-bpnn and SpAEbpnn in deep learning, and anomaly detection based StAE with RF and SpAE with RF which is selected as better prediction rate than SVM model. In this research, the assessment of proposed models was based on well-established area under the receiver operating characteristic curve. Parameter tuning was also utilized to assess better accuracy. Finally, landslide susceptibility maps were generated using equal interval function in ArcGIS 10.6 software.

Frequency ratio (FR) is determined by assessing the count of landslide pixel grids within each category, and the frequency ratio for each factor category is calculated by dividing the landslide ratio by the area ratio. The frequency ratio illustrates correlation between landslides and the influencing factors in a certain region. If this ratio is exceeds 1, then the relationship between a landslide and the affecting factor's class will be strong and conversely if the ratio is less than 1, then the relationship will be weak. If the value is 1, it means an average correlation (Meten et al. 2015). Support vector machine (SVM) Two main principles of SVM are the optimal classification hyperplane and the use of kernel features. The purpose of optimal sorting hyperplanes is to accurately distinguish between different classes of data in high-dimensional spaces.

Table.3.1. Exploratory Data Analysis

Unnam ed: 0	Te mp Hig h F	Te mp Avg F	Te mp Low F	Dew Point AvgF	Humi dityH igh Perce nt	Humi dityA vg Perce nt	Humi dityL ow Perce nt	Sea Level Pressure Avg Inches	Visib ility Avg Miles	Wind High MPH	Wind Avg MPH	Wind Gust MPH
1	57	49	40	37	94	69	44	31.13	10	17	7	26
2	59	46	33	28	77	53	28	31.49	10	9	4	13
3	62	47	32	29	90	57	23	31.45	10	13	5	21
4	59	51	42	41	87	72	57	31.33	10	11	3	17
5	58	49	40	37	80	64	48	31.4	9	13	4	18
6	61	54	46	40	84	66	48	31.39	9	8	2	12
7	63	52	41	40	93	65	37	31.17	10	11	3	15
8	65	51	37	42	93	77	61	31.1	10	18	6	25
9	45	41	36	27	76	61	46	31.33	10	14	6	22
10	56	47	37	29	77	55	33	31.39	10	9	2	13

Table.3.2. Display the last 10 rows of the data frame df_rainfall. Tail (10)

Temp High F	Temp Avg F	Dew Point Avg F	Humidity High Percent	Humidity Avg Percent	Humidity Low Percent	Sea Level Pressure Avg Inches	Visibility High Miles	Visibility Low Miles	Wind Avg MPH	Wind Gust MPH	Precipitation Sum Inches
1305	100	77	71	89	62	34	31	10	15	6	21
1307	102	78	71	89	61	32	30.1	10	19	6	30
1308	104	78	70	86	57	28	31	10	15	6	23
1309	105	78	69	86	57	28	31	10	15	6	24
1312	103	77	72	92	61	30	31	10	14	4	20
1313	104	79	72	92	62	32	31	10	16	8	25
1314	105	79	71	92	59	26	31	10	16	7	28
1315	104	76	68	83	55	26	31	10	13	6	22
1318	107	80	69	70	49	28	30.9	10	14	5	21
1319	100	78	70	65	44	23	31	10	13	5	21

4. Results and Discussions.

4.1. Real-time rainfall challenges forecasting and prediction with machine learning.

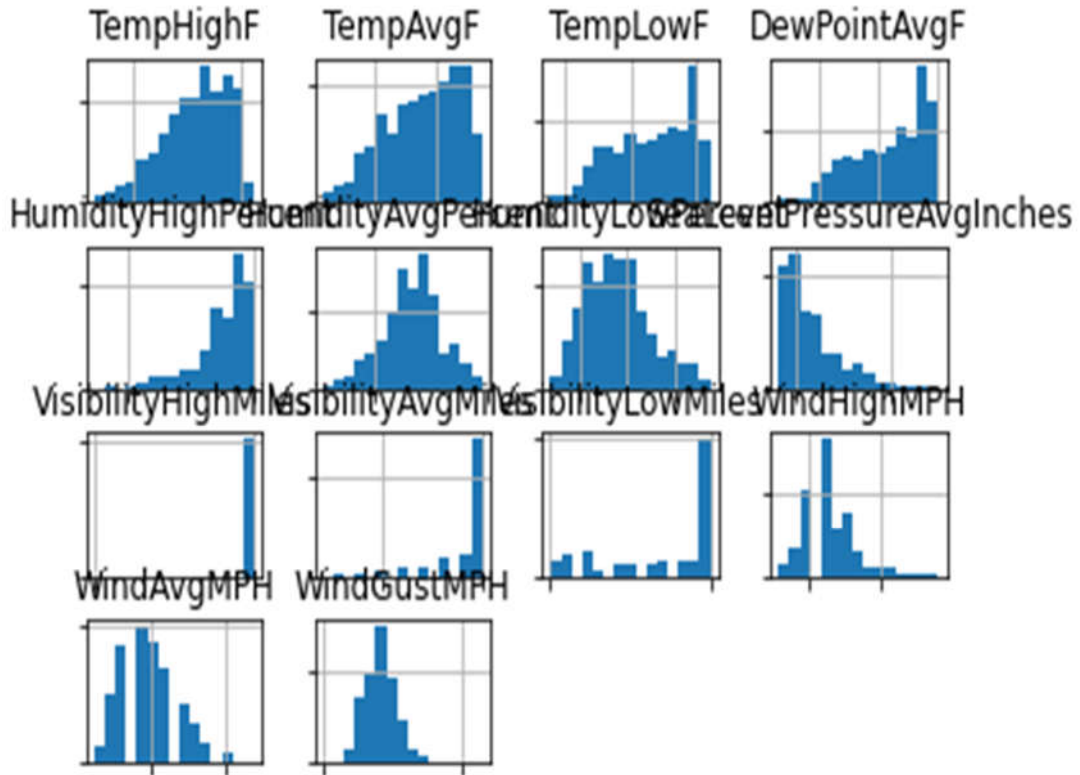


Fig.2 Temp High f, Dew point Temperature, humidity.

Based on historical weather data, such as temperature, dew point, and humidity as inputs to predict future events, a machine learning model predicts rain. A higher dew point and humidity, as well as a lower maximum temperature, are signs of rain. Deep learning models like LSTM capture complex, nonlinear relationships in the data better than models like Random Forest and Gradient Boosting. How ML models use these features for prediction. Several factors are taken into account in machine learning models to make predictions, including Temperature (both maximum and minimum), Dew Point Temperature, Humidity (Relative Humidity), Atmospheric Pressure, Wind speed and direction, and Cloud cover.

4. 2. Feature Importance of precipitation Amount.

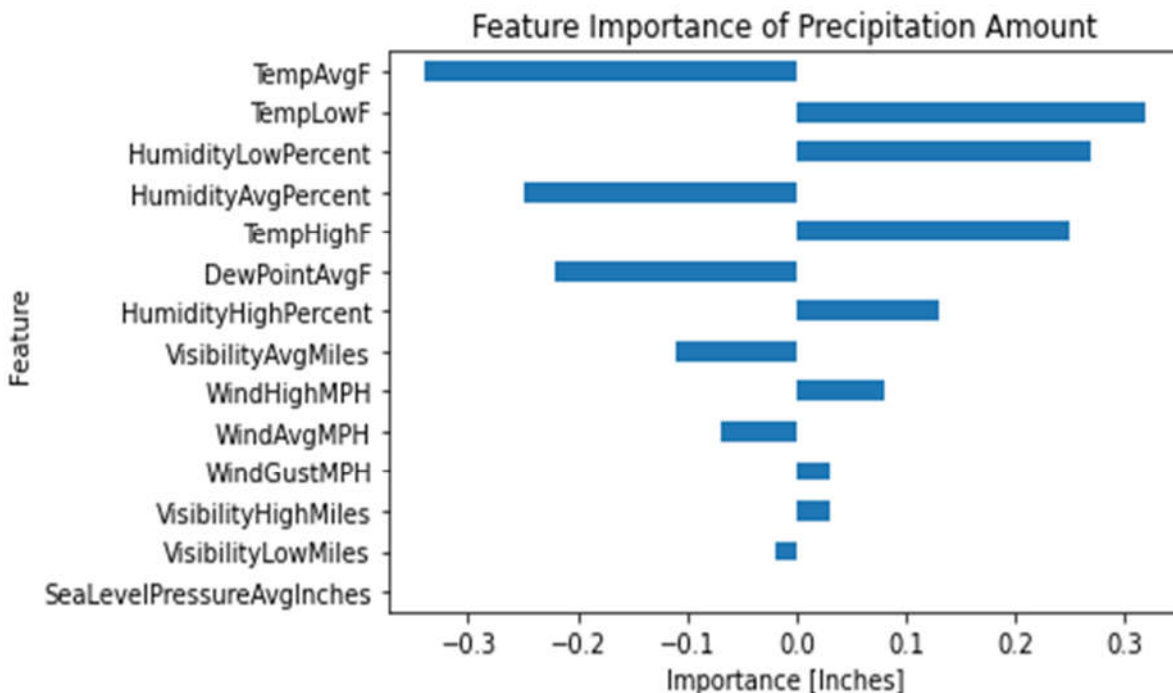


Fig.3 Feature Importance of precipitation Amount.

An important feature in rainfall prediction with machine learning is feature importance, which ranks input features according to how useful they are for predicting rainfall. The key features that are generally relate to atmospheric conditions that directly or indirectly impact precipitation. Relative humidity: The single most important feature frequency identified is Relative humidity.- More-water vapor in the atmosphere, or higher humidity is a direct cause of precipitation. Temperature is essential in determining the stability of atmosphere. The variation between the highest and lowest daily temperatures (diurnal variation) is also a significant factor.

Wind speed and direction: Wind is a key factor influencing weather patterns.

Wind speed and gust speed play a significant role, reflecting the motion of air masses that carry moisture. The direction of the wind influences the transport of that moisture. Variations in pressure systems are fundamental to predict weather. Low pressure systems frequently associated with development of cloud formation and precipitation. Sunshine duration: Reduced sunshine, signifying increased cloud cover, is often directly correlated with a higher chance of rain. Cloud cover: Serving as a clear visual sign of atmospheric humidity and conditions, cloud cover is among the most indicative characteristics. The rate of evaporation directly influences the amount of moisture present in the atmosphere.

4.2. Feature Importance of precipitation Amount.

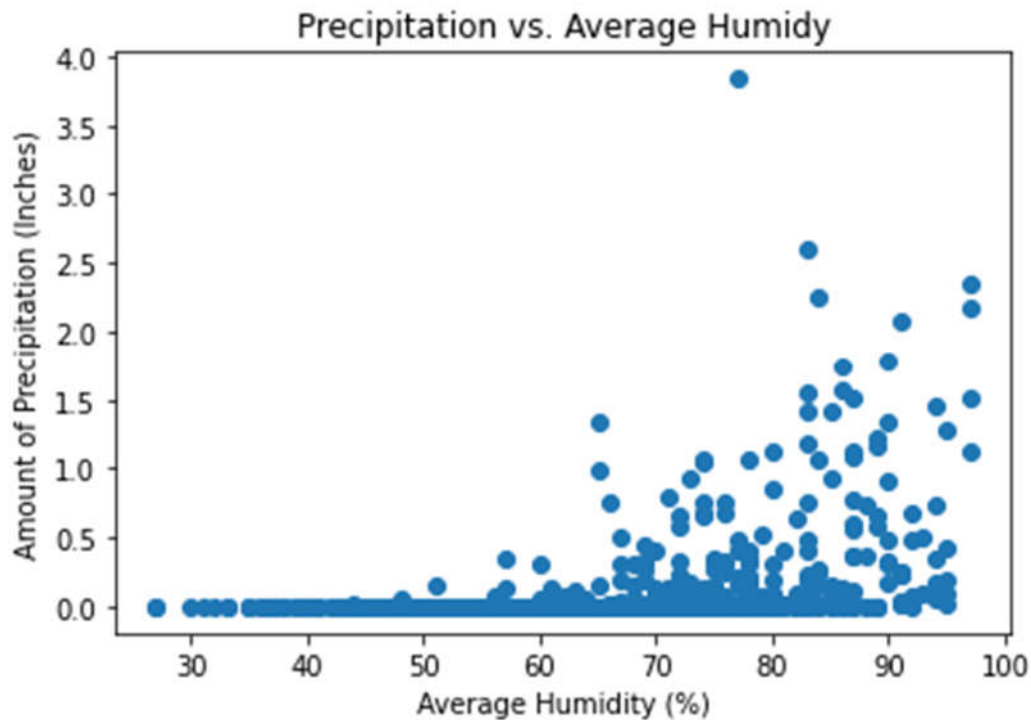


Fig.5 precipitation v/s Average humidity

The term "Fig: forecasting precipitation and average humidity drop using machine learning" denotes a figure or diagram found in a research paper or project report. This illustration visually depicts the outcomes or architecture of a machine learning model developed to predict variations in precipitation and average humidity. The figure would show the results of a machine learning experiment, probably featuring: A comparison between observed and anticipated values. A line or scatter graph could illustrate the degree to which the model's forecasts for rainfall and humidity correspond with actual, recorded data. Performance Metrics: Bar charts may illustrate the precision and error rates (e.g., Mean Absolute Error or Root Mean Squared Error) of various machine learning algorithms like, such as Random Forest or XGBoost when forecasting precipitation and humidity. The structure of the model: For deep learning methods, a diagram in flowchart format might illustrate the layers and organization of the neural network employed for forecasting the structure of the model.

4.4. Predicted Value v/s Actual Value

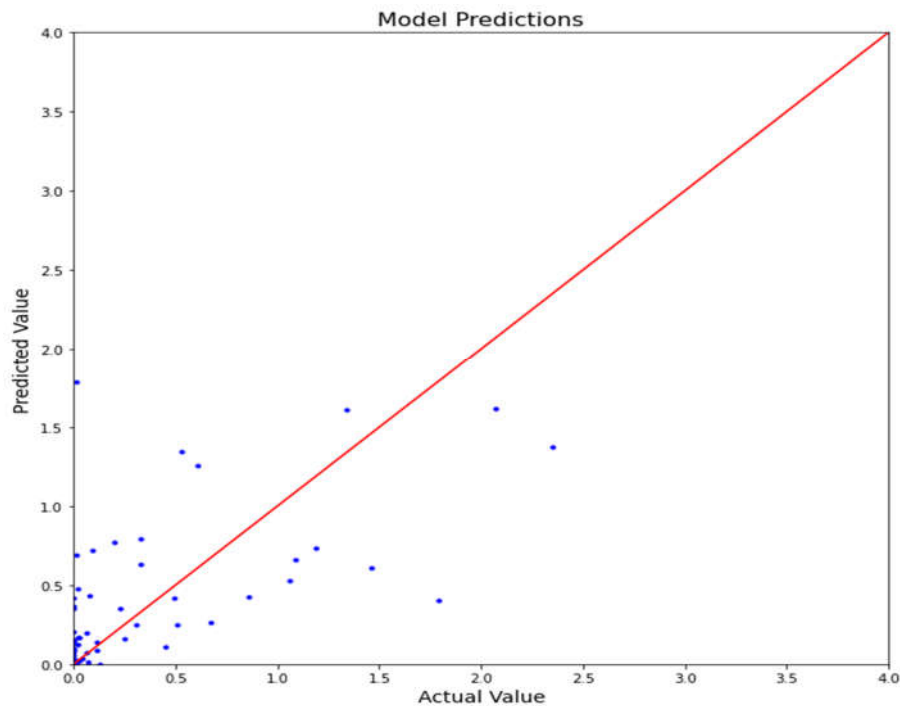


Fig.6 Predicted Value v/s Actual Value

Linear regression is an essential technique in predictive that uses historical data to predict an output variable. It's popular in predictive modeling for predictive modelling because it is easily understood and the capacity to articulate concepts in simple terms. Linear regression models are applied in diverse real-world scenarios multiple sectors like economics (for example, forecasting growth), business (such as estimating sales of product or assessing employee performance), social science (like predicting political preferences influenced by gender or race), healthcare (for instance, anticipating blood pressure levels based on weight or the onset of diseases from biological aspects), and more. Understanding to apply linear regression models can uncover insights within data to address important issues. Figure 4.6 illustrates the comparison of correlation coefficients between observed and forecasted rainfall for a 1-month lead time at various stations throughout the validation period from 2009 to 2018. Based on figure 4.5, it is clear that the stochastic efficiency is less than input selection at rainfall stations. This indicates that the suggested approach using DNN with an input selection technique yields superior results compared to DNN model technique lacking an input selection technique. Applying the selection method, the anticipated and documented displays reduced variability in contrast to not using the input selection method.

4.5. Coefficient magnitude and coefficient index.

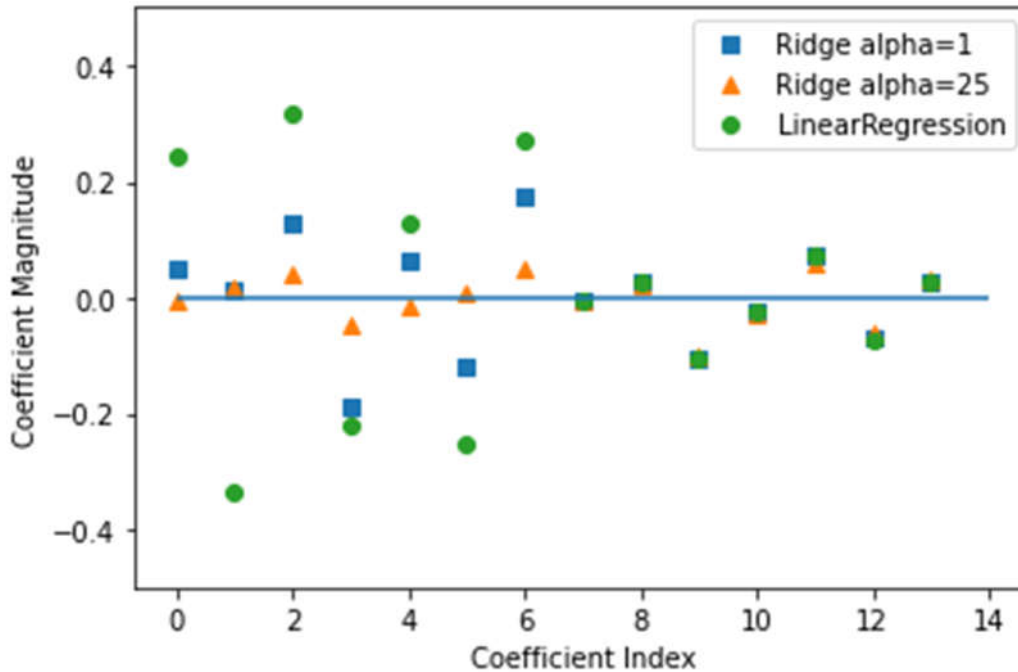


Fig.7 Coefficient magnitude and coefficient index.

In rainfall prediction using machine learning, a "coefficient magnitude" chart shows the significance of different weather-related input features. In rainfall prediction using machine learning, a "coefficient magnitude" chart shows the significance of different weather-related input features. In the absence of specific data from Fig. 7, a general interpretation of a regression model involve. Coefficient magnitude (y-axis): The absolute value of the coefficient reflects the strength of a feature impact on the prediction. A higher magnitude signifies greater importance

- Coefficient index (x-axis): This represents a numbered or indexed label assigned to each input feature used in the model, such as humidity, wind speed, or temperature. Consider a streamlined rainfall prediction model that operates using three primary inputs: relative humidity, wind speed, and sea level pressure. The associated graph would display three indices along the x-axis indicated as 1, 2, and 3.

4.6. Daily Rain Fall Report

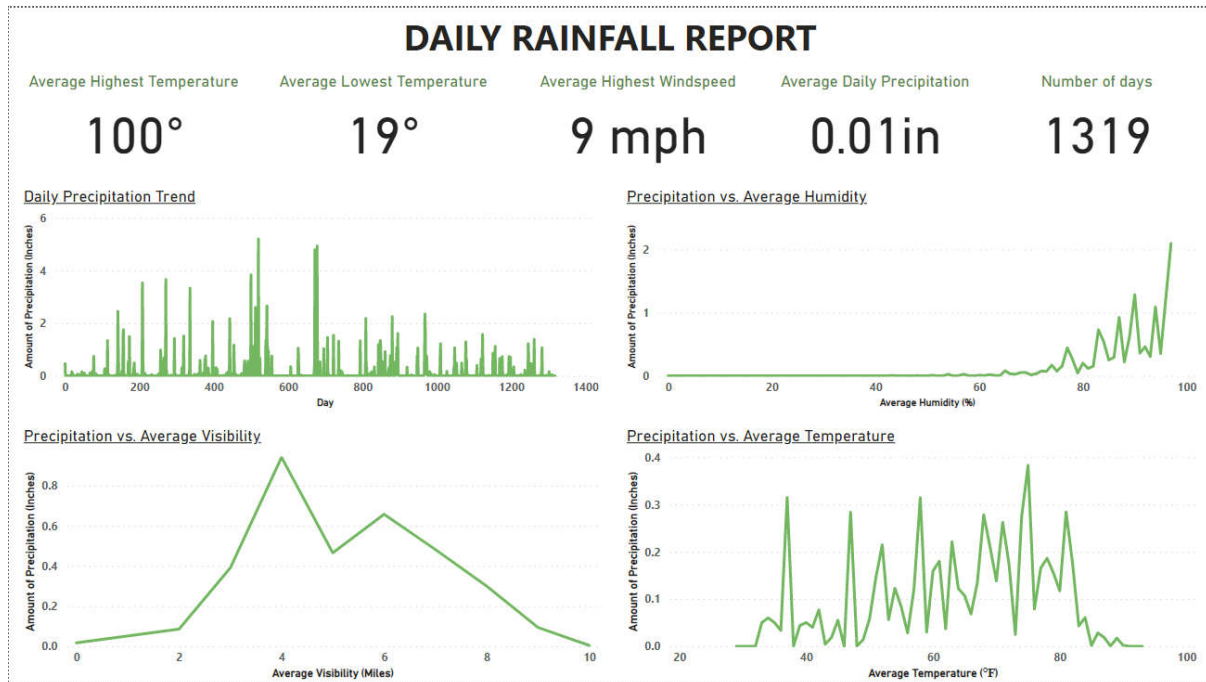


Fig:8 daily rain fall report

An efficient daily rainfall report generated through machine learning depends on using historical meteorological data to forecast future precipitation. The process entails collecting and processing an extensive dataset, developing and training predictive models, and assessing their accuracy to produce an informative report.

4.7. Data collection and feature engineering

The model uses historical data from meteorological stations to analyse patterns and relationships among various weather variables. Important features include atmospheric conditions: such as temperatures maximum and minimum temperatures, humidity levels, and sea-level pressure which are crucial indicators of rainfall. For example, elevated frequently signals an approaching rainstorm.

Wind data: Wind speed and direction significantly impact the movement of weather systems.
 Geographic information: Local topography and geographical location are crucial in shaping rainfall patterns.
 Time-series features: The time of year, including the month and season, plays an essential role in predicting rainfall.

4.8. Distribution of highest Temperature

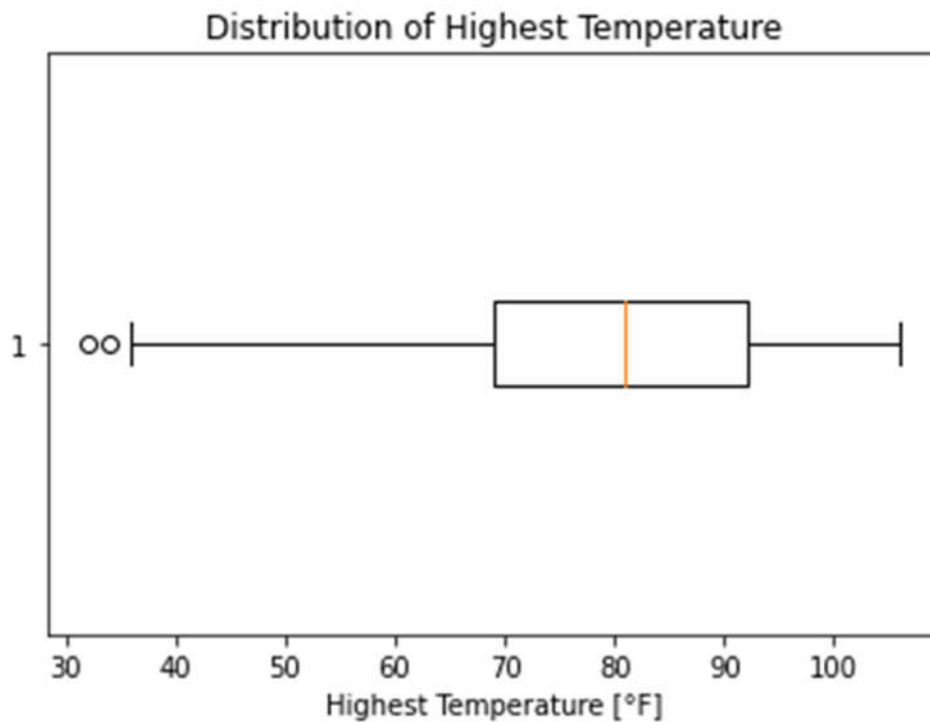


Fig.9 Distribution of highest Temperature

To forecast the distribution of highest temperature and rainfall through machine learning, researchers generally adopt one of two main approaches: probabilistic modeling or quantile regression. These methodologies utilize incorporate historical meteorological data to forecast future weather patterns while addressing the inherent uncertainties present in weather systems.

4.9. Methods for modeling probability distributions

These models go beyond predicting individual values by forecasting the entire probability distribution of a variable, approach is essential for anticipating extreme events and facilitates risk evaluation, such as estimating the chances of temperatures surpassing a specific limit or the probability of intense rainfall.

4.10. Mixture Density Networks (MDNs)

MDNs integrate a traditional neural network such as a combination

1. a mixture of Gaussian distributions).

- How it works: The neural network output from an MDN determines parameters of a mixture distribution, encompassing the mean, variance, and weight for each component. This enables the model to represent multi-modal distributions effectively, such as the distinct temperature patterns observed during dry and wet seasons. Example use: application: Estimating the complete probability distribution of maximum daily rainfall, with the output represented as a set of parameters combination of Weibull or Gamma distributions.

2. Probabilistic weather forecasting with deep learning

Advanced data-intensive weather forecasting systems such as those employed by the European Centre for Medium-Range Weather Forecasts (ECMWF) use deep learning models like Graph Neural Networks (GNNs) and Transformers. These sophisticated models enable the prediction of probability density functions for weather variables, effectively representing the complete spectrum of potential future scenarios.

4.11. Distribution of highest Temperature

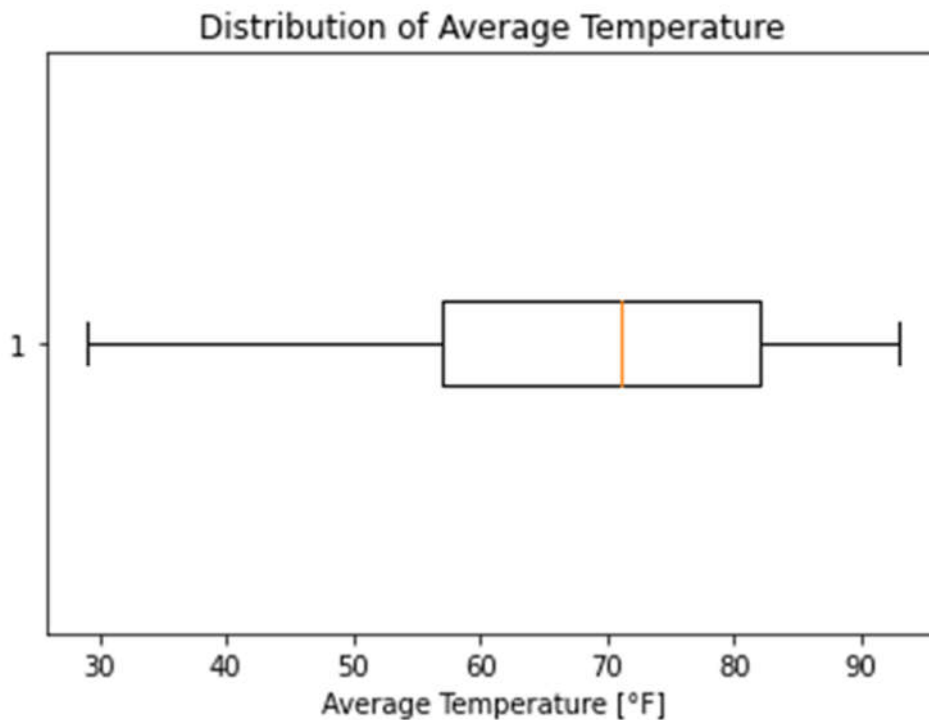


Fig: 10 Distribution of highest Temperature

A machine learning-based rainfall prediction system determines patterns linking atmospheric factors such as elevated temperatures to the probability of rain. In this framework, the variation of high temperatures serves as a crucial input variable. By processing historical data, the model evaluates how the occurrence and severity of high temperatures impact weather scenarios that generally lead to rainfall.

4.12. How machine learning uses temperature distribution to predict rainfall.

Data collection and feature selection

The process starts with collecting historical meteorological data from multiple sources, including weather stations and satellites. Encompasses several key variables: Maximum Temperature representing the highest temperature recorded on a specific day: Humidity indicating the moisture content in the air. Dew Point which marks the temperature at which the air becomes saturated with water vapour: Atmospheric Pressure denoting the force exerted by the atmosphere, Cloud Cover measuring the portion of the sky obscured by clouds and Wind Speed and Direction describing the movement of air. .

4.13. Analyzing temperature distribution

The relationship between the highest temperature distribution and rainfall is examined alongside other characteristics. Exploratory data analysis (EDA) and statistical tests are utilized to uncover significant patterns. For instance, research frequently highlights that elevated temperatures may act as a precursor to rainfall, particularly in humid environments. Higher temperatures can accelerate evaporation, introducing more moisture into the atmosphere and consequently boosting the likelihood of precipitation.

5. Relationship between the dependent variables and the independent variable

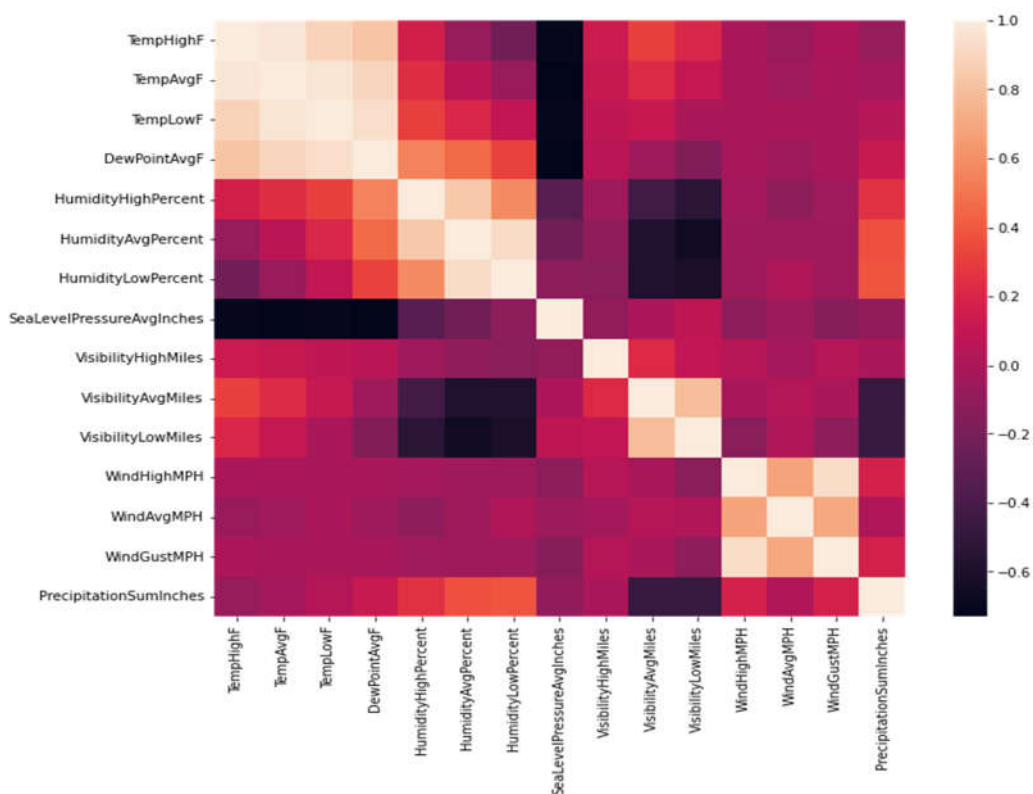


Fig. Relationship between the dependent variables and the independent variable

In machine learning for rainfall prediction, the the model determines the relationship by analyzing how independent variables such as historical weather data correlate with and affect the dependent variable which is rainfall. The model examines intricate patterns and relationships to generate predictions for the dependent variable when presented with new data. This dependent variable represents the target or outcome that the machine learning model is designed to forecast.

In this context, it is rainfall, which can be defined using different approaches. Binary classification: Determining whether it will rain or not. Continuous regression: Estimating the precise amount or intensity of rainfall such as measurements in millimetres. The independent variables would typically include weather and atmospheric data. Independent variables often referred to as features or predictors, are the inputs used in training a machine learning model. These variables represent factors affect rainfall either directly or indirectly. Common independent variables include: Temperature—Represented by the maximum and minimum daily measurements Humidity – Indicates the concentration of water vapour in the air. Air pressure – refers to variations in atmospheric pressure that, influence weather patterns. Wind speed and direction influence the movement of weather systems and the distribution of moisture. Cloud cover refers to the portion of the sky obscured by clouds. Evaporation measures how quickly moisture transitions into the atmosphere. Geographical factors, such as location, elevation, and proximity to water bodies also play a significant role. Lastly, solar radiation, the energy from the sun, drives both evaporation processes and temperature variations.

6. Conclusion

Weather forecasting involves meteorological processes that enable researchers to refine their work through the application of numerical weather prediction techniques. Weather prediction is carried out by using various data mining techniques including classification clustering and neural network, decision trees. This study focuses on enhancing the classification and prediction performance within traditional frameworks. A weather prediction model has been developed as part of the research; however, certain limitations of the model have been identified, indicating the need for a thorough review in near future need to be review before implementing the proposed techniques. Furthermore, challenges remain in optimizing the application of these data mining methods in the field of weather forecasting, highlighting areas for improvement in future implementations.

7. References

1. Latif, S. D., Hazrin, N. A. B., Koo, C. H., Ng, J. L., Chaplot, B., Huang, Y. F., ... & Ahmed, A. N. (2023). Assessing rainfall prediction models: Exploring the advantages of machine learning and remote sensing approaches. *Alexandria Engineering Journal*, 82, 16-25. <https://doi.org/10.1016/j.aej.2023.09.060>
2. Kandasamy, O., N, M., E, S., R, R., Kannan, B., & C, P. (2025). Rainfall prediction using artificial neural networks and machine learning algorithms over the Coimbatore region. *Journal of Water and Climate Change*, 16(3), 946-959. <https://doi.org/10.2166/wcc.2025.576>
3. Basha, C. Z., Bhavana, N., & Bhavya, P. (2020, July). Rainfall prediction using machine learning & deep learning techniques. In *2020 international conference on electronics and sustainable communication systems (ICESC)* (pp. 92-97). IEEE.
 - a. DOI: [10.1109/ICESC48915.2020.9155896](https://doi.org/10.1109/ICESC48915.2020.9155896)
4. Kandasamy, O., N, M., E, S., R, R., Kannan, B., & C, P. (2025). Rainfall prediction using artificial neural networks and machine learning algorithms over the Coimbatore region. *Journal of Water and Climate Change*, 16(3), 946-959. <https://doi.org/10.2166/wcc.2025.576>.
5. Preethi, B. M., Gowtham, R., Aishvarya, S., Karthick, S., & Sabareesh, D. G. (2021). Rainfall prediction using machine learning and deep learning algorithms. *International Journal of Recent Technology and Engineering (IJRTE)*, 10(4), 251-254.
6. Nam, K., Wang, F. An extreme rainfall-induced landslide susceptibility assessment using autoencoder combined with random forest in Shimane Prefecture, Japan. *Geoenviron Disasters* 7, 6 (2020). <https://doi.org/10.1186/s40677-020-0143-7>
7. Rahman, A. U., Abbas, S., Gollapalli, M., Ahmed, R., Aftab, S., Ahmad, M., ... & Mosavi, A. (2022). Rainfall prediction system using machine learning fusion for smart cities. *Sensors*, 22(9), 3504. <https://doi.org/10.3390/s22093504>.
8. Yue, T., Zhang, S., Zhang, J., Zhang, B., & Li, R. (2020). Variation of representative rainfall time series length for rainwater harvesting modelling in different climatic zones. *Journal of Environmental Management*, 269, 110731. <https://doi.org/10.1016/j.jenvman.2020.110731>.
9. S. Rani, A.K. Mishra, A. Kataria, S. Mallik, H. Qin, Machine learning-based optimal crop selection system in smart agriculture. *Sci. Rep.* 13(1), 15997 (2023). <https://doi.org/10.1038/s41598-023-42356-y>

10. S. Manandhar, S. Dev, Y.H. Lee, Y.S. Meng, S. Winkler, A Data-Driven Approach for Accurate Rainfall Prediction, *IEEE Trans. Geosci. Remote Sens.* 57 (2019) 9323–9330, <https://doi.org/10.1109/TGRS.2019.2926110>.
11. M.A. Boucher, D. Tremblay, L. Delorme, L. Perreault, F. Anctil, Hydro-economic assessment of hydrological forecasting systems, *J. Hydrol.* 416–417 (2012) 133–144, <https://doi.org/10.1016/j.jhydrol.2011.11.042>.
12. M.C. Urban, G. Bocedi, A.P. Hendry, J.B. Mihoub, G. Pe'er, A. Singer, J.R. Bridle, L. G. Crozier, L. De Meester, W. Godsoe, A. Gonzalez, J.J. Hellmann, R.D. Holt, A. Huth, K. Johst, C.B. Krug, P.W. Leadley, S.C.F. Palmer, J.H. Pantel, A. Schmitz, P.A. Zollner, J.M.J. Travis,, Improving the forecast for biodiversity under climate change, *Science* 80- (2016) 353, <https://doi.org/10.1126/science.aad8466>.
13. P. Umamaheswari, V. Ramaswamy, Optimized Preprocessing using Time Variant Particle Swarm Optimization (TVPSO) and Deep Learning on Rainfall Data, *J. Sci. Ind. Res.* 81 (2022), <https://doi.org/10.56042/jsir.v81i12.69310>.
14. S. Salma, A. Ashwitha, Hybrid CNN-LSTM Model: Rainfall Analysis and Prediction for Karnataka Region, *J. Theor. Appl. Inf. Technol.* 100 (2022).
15. S. Iliger, M. Pinto, Comparative Study of Rainfall Prediction Modeling Techniques: Case Study on Karapur, India, *Artif. Intell. Commun, Technol*, 2022 <https://doi.org/10.52458/978-81-955020-5-9-16>.
16. D. Wu, L. Wu, T. Zhang, W. Zhang, J. Huang, X. Wang, Short-Term Rainfall Prediction Based on Radar Echo Using an Improved Self-Attention PredRNN Deep Learning Model, *Atmosphere (basel)*. (2022), <https://doi.org/10.3390/atmos13121963>.
17. R. Fredyan, G.P. Kusuma, Spatiotemporal Convolutional Lstm With Attention Mechanism for Monthly Rainfall Prediction, *Commun. Math. Biol. Neurosci.* 2022 (2022). <https://doi.org/10.28919/cmbn/7761>.
18. I.N. Ikpang, E.J. Okon, E.U. George, Modeling Average Rainfall in Nigeria With Artificial Neural Network (ANN) Models and Seasonal Autoregressive Integrated Moving Average (SARIMA) Models, *Int. J. Stat. Probab.* 13 (2022), <https://doi.org/10.5539/ijsp.v11n4p53>.
19. I.R. Ilaboya, S.N. Nwachukwu, Development of Intensity Duration Frequency (Idf) Curves for Rainfall Prediction in Some Selected States in South-West Nigeria, *J. Energy Technol. Environ.* 4 (2022). <https://ccsenet.org/journal/index.php/ijsp/article/view/0/47422>.
20. D.M.S. Zekrifa, M. Kulkarni, A. Bhagyalakshmi, N. Devireddy, S. Gupta, S. Boopathi, Integrating machine learning and ai for improved hydrological modeling and water resource management. In: *Artificial intelligence applications in water treatment and water resource management*, pp. 46–70. IGI Global (2023). <https://doi.org/10.4018/978-1-6684-6791-6.ch003>
