IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14123

RAINFALL PREDICTION AND AGRICULTURE ANALYSIS USING MACHINE LEARNING

Mr. Sreenivasa M[1], Abhishek R[2], Akash Hosur[3], Battula Shabareesh[4], Tarun B[5]

Assistant Professor, Ballari Institute of Technology and Management, India[1]

Students, Computer Science Engineering, Ballari Institute of Technology and Management, India[2-5]

Abstract: Rainfall forecasting plays an important role in increasing agricultural production and decreasing associated risks resulting from climate change. This is especially because traditional approaches to rainfall forecasting often do not adequately capture the complex and nonlinear features of climate change. Hence, this research examines the applicability of using machine learning algorithms, such as decision tree regressors, for the yearly and monthly rain events based on the historical climate and geographical information. Apart from incorporating agricultural analysis to advise potential appropriate crops in accordance with the expected amount of rainfall, type of soil, as well as its pH level, the proposed application also brings experimental results affording better tools for decision-making for farmers and other stakeholders in the agricultural field. This study illustrates the possibility of applying quantitative approaches to promote sustainable farming practices and, therefore, achieve food security in conditions when weather patterns become volatile.

Keywords: Rainfall Prediction, Machine Learning, Decision Tree Regressor, Agricultural Analysis, Crop Recommendation.

I. INTRODUCTION

Within the climatic conditions, there is none as influential as the rainfall as regards production outputs of an agricultural nature. Rainfall has its importance in farming and crop production, irrigation planning and scheduling, and crop and farm management. In the areas where the rainfall is high or unpredictable, then an accurate rainfall prediction can decrease the effects of drought, floods, and other difficult situations due to climate change. Many of the statistical models as well as observational methods used to predict rainfall have shortcomings in representing and analyzing the complex aspects of the convective systems, owing to the non-linearity and coupling aspects in climatic data. It is from this argument that scholars' interest in incorporating ML models into systems for accurate rainfall prediction has gained momentum.

Machine learning techniques also then hold great potential for a holistic approach to solving problems in the extension of the prediction of phenomena, including weather and climatic determination. This is distinct from conventional statistical techniques because the focused, created ML algorithms do not have any problem in handling large, multifaceted data, discovering latent patterns, and making changes to new underlying climatic factors. On that basis, it is possible for accurate machine learning models to predict rain in a given locality but also inform adaptive agricultural practices. This brings improvement of the reliability and usability of the rainfall prediction in agriculture needed by farmers, planners, and policy makers.

In this study, therefore, our analysis centered on machine learning models in rainfall prediction, especially decision tree regressors, a frequently used model that features interpretability and the ability to capture interactions between predictor variables. The goal is to forecast by year and month and rainfall by using the climatology from meteorologic data and geography about districts, taluks, and hoblis (subdivisions of districts in regions of India). Rainfall data is accumulated for several years and encompasses not only the overall annual measurement but also the measurement for each particular month, which gives additional clues to the seasonal fluctuations in rainfall.

Rainfall forecasting has a measurable effect on the choice of crops to be grown and overall yields in agriculture. With the present rising concern in climate change, its erratic nature and changes in trends, this kind of information will go a long way in helping farmers to make decisions on dates to sow, when to irrigate, and which crops to cultivate. In addition, this enables the farmers to implement measures of risk minimization. The various hazards, including possible drought and excessive rainfall, are damaging to crop cultivation. Therefore, the integration of rainfall prediction models into agricultural analysis is important in formulating sustainable farming methodologies.
Apart from the rainfall estimation, this work also has an agricultural estimator part.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14123

Crops that need to be grown in the areas estimated before the rain, the type of soil in that area, as well as the soil pH in those areas. There is variation in the water needs and the ability to grow in a particular subsoil. For instance, rainfall conditions favored rice, while conditions that favored crops like sorghum or millet are conditions that are normally received as relatively low rainfall. This agri-allometry may assist farmers in getting crop recommendations likely to optimize crop yields and use of resources and land appropriately.

The Decision Tree Regressor, for example, can even Identify the key drivers that affect rainfall and crop suitability. This would have benefited in establishing the trend of the regional climate. The model does not only forecast mean rainfall for the district, individual taluk, and hobli, but also the change in this parameter in the course of the month is crucial for accurate irrigation and crop growing. Consequently, it addresses the gap between the forecast for rainfall and actual decision-making during agriculture with climatic uncertainty through its integration and database strategy.

II. STUDY AREA AND DATA SOURCE

A. STUDY AREA

The area of focus of this research is Karnataka state in India that experience a variety of climatic conditions from the arid to tropical. Karnataka has a major chunk of its agricultural land in the region and agrarian operations greatly depend on monsoon rains; hence the region can be established as most important for studying rainfall prediction and analysis in India. The state is divided into numerous districts, and this study focuses on two districts: Ballari and Dharwad.

Ballari District: Located in the northern part of Karnataka, Ballari is predominantly an arid to semi-arid region. This district has, for all intents and purposes, had a hot summer with only moderate showers during the monsoon. One of the major economic activities done in this district is farming; this includes groundnuts, sunflowers, cotton, and sorghums. Yield fluctuations are in fact attributed almost wholly to fluctuations in rainfall, especially uneven distribution of rainfall; accurate prediction of rainfall is therefore very vital to optimize agricultural inputs productivity as well as to more efficient use of water resources.

Dharwad District: Dharwad is located at the heart of Karnataka. The climate is relatively not that heavy, but raining is much heavier during the monsoons than at Ballari. Some of the agricultural activities carried out in this district include the farming of rice, sugarcane, cotton, and tomatoes. Here again, relatively larger variations in empirical rainfall make this place a more interesting place to study rainfall and the impact of differential rainfall on different types of crops in this climate condition as compared to Ballari. prediction for both the semi-arid and moderate rainfall regions and the implications of these rains to crop choices and district, taluk, and hobli levels of agricultural activities.

B. DATA SOURCE

1. **Rainfall Data**: The IMD mainly generates historical rainfall data for various districts in Karnataka and has been relied upon in this study. The records found in this dataset comprise accurate monthly rainfall data of the districts of Ballari and Dharwad for some years. This data makes it possible to define the temporal periodicity of the rainfall and its annual fluctuations and, thus, contribute to more successful agricultural practices.

Rainfall data is also provided in monthly and annual terms. For instance, the rainfall is comparatively lesser for Ballari than Dharwad – and that amounts to the climate disparity between the two districts.

Sample rainfall data for Ballari and Dharwad:

This table gives the Ballari and Dharwad districts' rainfall for each of the months as the average rainfall in mm calculated over several years. As we have analysis, Ballari receives less rainfall than Dharwad and Dharwad receives good amount of rainfall in the monsoon period means June to September.

District	Taluk	Hobli	Annual Rainfall (mm)	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Ballari	Hospet	Hospet 1	800	30	25	35	50	100	140	160	170	110	95	70	30
Ballari	Ballari	Ballari 1	850	40	30	40	60	110	150	175	180	120	100	75	35
Dharwad	Hubli	Hubli 1	1200	50	45	55	80	130	180	200	210	150	130	100	60
Dharwad	Dharwad	Dharwad 1	1250	60	50	60	90	140	190	220	230	160	140	110	65

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14123

2. **Agricultural Data**: Information regarding specific types of soil, suitable crops, or other climatic factors that control crop growing and production in such areas refers to agricultural information. The dataset, therefore, comes from agricultural institutions as well as state departments of agriculture and crop suitability reports. This information helps in knowing which crops need how much rainfall, soil type, and so on, for their successful cultivation in the Ballari and Dharwad areas.

For example, while Ballari receives somewhat less rain and less frequent rainfall than Dharwad, it is more suited to groundnuts and sunflowers, while heavy rain and frequent rainfall in Dharwad are suitable for paddy and sugarcane crops.

Example agricultural data snippet for crop suitability:
The crop is determined compatible with rainfall, soil, temperature, etc some characteristics like pH, fertilizer..etc. Students can perhaps know how the various crops are suited to the rainfall distribution in the two districts of Ballari and Dharwad.

Data Preprocessing
These data sets of collected information must be passed on to the system through various stages of preprocessing as follows:

Cleaning of Data: That means missing values are to be addressed, and if any kind of error is present in the precipitation data and agriculture data then outliers are required to be addressed.

Crop	Suitable Districts	Minimum Rainfall (mm)	Soil Type	Temperature Range	Suitable pH Range	Recommended Fertilizer
Groundnut	Ballari	400	Sandy Loam	20°C - 35°C	6.0 - 7.5	Nitrogen, Potassium
Sunflower	Ballari	500	Clay Loam	18°C - 30°C	6.0 - 7.5	Phosphorus, Nitrogen
Paddy	Dharwad	900	Clay	25°C - 32°C	5.5 - 7.0	Nitrogen, Potassium
Sugarcane	Dharwad	1000	Alluvial	20°C - 35°C	6.0 - 7.0	Phosphorus, Potassium
Cotton	Ballari, Dharwad	600	Sandy Loam	22°C - 30°C	6.5 - 7.5	Nitrogen, Potassium

Encoding: Categorical data, for example district, taluk, hobli, among others, will have to undergo the Label Encoding technique making the input of the machine learning algorithms numbers.
Normalization: Any variable that is in the continuous type is normalized including rainfall data. The normalization enables machine learning algorithms to find their best working condition.
Feature Engineering: Building meaningful features for the system and to increase it precision such as the monthly rainfall, monthly variations, and forecasted annual rainfall.

By employing the detailed data collected for Ballari and Dharwad, this study aims at forecasting the rainfall regimes while also offering crop recommendations in a manner that helps the farmers to understand the best ways to practice agriculture to maximize the yields as well as plan for water use adequately. One very important measure that can be derived through the use of machine learning techniques and is focused on this practical region-specific solution is that concerning climate change and variability in regards to the weather.

III. METHODOLOGY

The method for this study comprises the following steps: data acquisition, data cleaning and normalization, and modeling and prediction of rainfall and agriculture. This also uses machine learning algorithms to make rainfall forecasts based on rainfall data and suggest crops to be grown in a given weather and geographical location given specific soil type, amount of rainfall expected and probable temperatures.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14123

Steps Involved in the Methodology

1. Data Collection: The first one is data gathering from different authoritative sources:

- About rainfall data the information from the historical data information from the Indian Meteorological Department is used.
- Agricultural source tells about the crops' suitability for enterprise besides things like types of soils, similar other factors will be included in the agricultural departments or research institutes.
- The information source for geography department includes geographical data obtained from the Karnataka GIS.for the rainfall data includes the historical data information from the Indian Meteorological Department.
- Agricultural source details crop suitability along with other things like types of soils and similar more factors will include agricultural departments or research institutes.

2. Data preprocessing: Cleaning data is carried out in different formats that need a correct approach to format which should match with the targeted machine learning model.

- Cleaning Missing Values: With reference to the analysis of rainfall data, it means, excluding all rows or columns containing missing values.
- Label Encoding: Applying the label encoding technique, this categorical data is converted into numeric form in district, taluk, and hobli for preparing the data for modeling.
- Feature Engineering: New line features have been developed from the existing data; for example, monthly average, seasonal changes, and annual rainfall forecast.
- Normalization: The data like rain fall and temperature has been normalized and it has been done to make the data at around the right scale for better modelling.

3. Model Training and Validation: The approach applied most frequently during this research is Decision Tree Regressor and it fits for the continuous variables prediction, including rainfall. Hence train_test_split function is applied for historical rainfall data to divide in training and test datasets: for predictions of continuous variables such as rainfall:

- The dataset is divided between training and test sets by applying the train_test_split function for historical rainfall. Then apply the Decision Tree Regressor function with the training set.By the use of the
- The Model's Performance : Employ performance attributes, which explain its level of accuracy. In other words, it measures the distance in between predictions and the real values by the use of statistical techniques such as Mean Absolute Error, Root Mean Square Error, coefficient of determination to show how well the model emulates and predicts rainfall.

4. Rainfall Prediction: We will get an annual and a monthly predictive analysis on the rainfall for several districts, taluks, and hoblis once the model has been trained. Lastly, the input features which have been passed through the trained model are: district, taluk, and hobli for the predicted rainfall.

5. Agricultural Analysis and Crop Recommendation: Taking the predicted rainfall value into consideration and the soil condition on the field, crop suitability will be decided by the model.

- Soil pH and Type: Agricultural analysis contains user parameters such as the pH of the soil and type of the soil to define the crop to grow.
- Rainfall Analysis: The model based on the values of forecasted rainfall and historical data also suggests the necessary crops best suited for the climate and the ground of the mentioned region.
- Crop Recommendation: The model only allows the crops that would grow well in the soil, based on the PH and expected rainfall and the user input, the model suggests good crops to plant such as groundnuts, cotton, or paddy.

6. Visualization: To facilitate easy comprehension and comparison, the results are described together with illustrations:

- Pie and Bar Charts: Monthly rainfall distributions should be displayed to show readers a picture of how rainfall is likely to be distributed in a given month.
- Crop Suitability Graphs: A comparison of various crops in relation to rainfall, types of soils and their respective pH levels.

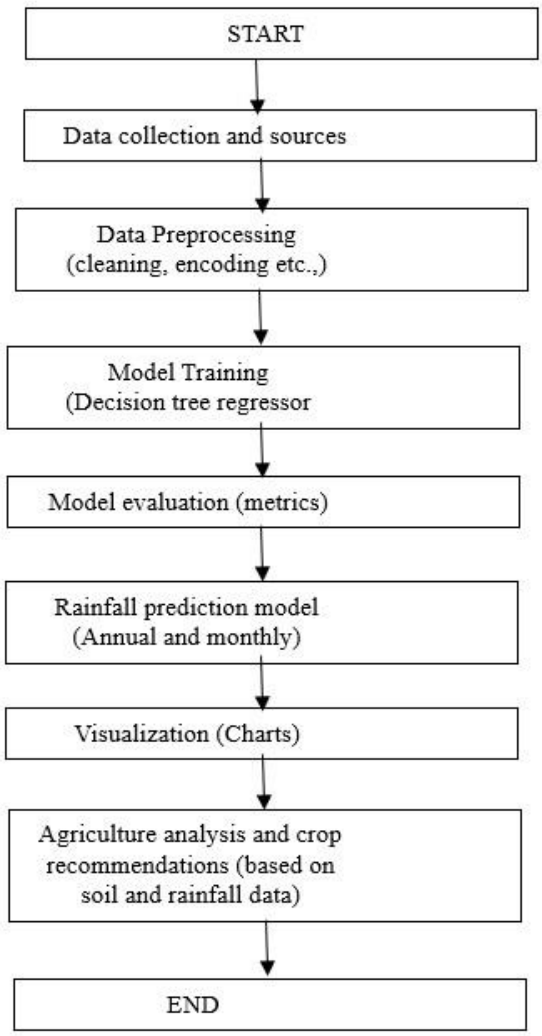
Flowchart Explanation:

1. Data Collection and Source: This are data gathering techniques that involve the use of rainfall, agricultural and geographical information.
2. Data Preprocessing: Therefore some of the preprocessing steps employed in the present work include data cleaning, label encoding of categorical variables and creating feature for better input into the model.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 14, Issue 1, January 2025
DOI: 10.17148/IJARCCE.2025.14123

3. Model Training: It is important now to split the data up into two sets: training data set and the data to test the accuracy of the Decision Tree Regressor on.
4. Model Evaluation: Evaluation metrics are expressed in Mean Absolute Error, Root Mean Squared Error and coefficient of determination R-squared.
5. Rainfall Prediction: In this case with the use of the trained model an approximation of annual and monthly rainfall for the various regions is obtained.
6. Agricultural Analysis and Crop Recommendation: Rainfall and conditions of the soil employed in the technologies are described as recommend crops for the areas of interest.

Visualization: This type of outputs are presented in a form of a picture such as a pie chart and a bar chart in which users are well capable of decoding properly the information that is provided. Flowchart of the methodology:



Visulizations of Methodology
The following are examples of types of visualizations used to explain the results:
Rainfall Distribution (Bar Chart):
Perhaps a bar graph of the predicted levels of rainfall in a district, let’s say Ballari or Dharwad, would be useful towards explaining how the rainfall is distributed in a particular year in that region.
This is where an actual chart, or what is here referred to as the output, should be.
Monthly distribution of rainfall – Pie Chart
This will simply be a pie chart with contribution by rainfall each month which helps to determine the peak seasons.
The chart below shows real values on the graph (actual chart(output)).

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025
DOI: 10.17148/IJARCCE.2025.14123

MODELS

The formulas that we are going to use to find the efficiency of each models are as follows:
1. MAE (Mean Absolute Error):
MAE = $1/n * \Sigma |y_i - \hat{y}_i|$
Where:
- y_i = Actual value of the i-th data point
- $\hat{y}_i$ = Predicted value of the i-th data point
- n = Total number of data points

2. RMSE (Root Mean Squared Error):
RMSE = $\sqrt{[(1/n) * \Sigma (y_i - \hat{y}_i)^2]}$
Where:
- y_i = Actual value of the i-th data point
- $\hat{y}_i$ = Predicted value of the i-th data point
- n = Total number of data points

3. R-squared:
$R^2 = 1 - [\Sigma (y_i - \hat{y}_i)^2 / \Sigma (y_i - \bar{y})^2]$
Where:
- y_i = Actual value of the i-th data point
- $\hat{y}_i$ = Predicted value of the i-th data point
- $\bar{y}$ = Mean of the actual values
- n = Total number of data points

1. **Decision Tree Regressor**
Decision Tree Regressor creates a decision tree where some purposes are error minimization, for example Mean Squared Error of samples that reach a split node.
It is good when There is no linear correlation between the independent or any variables, interpretable model However it biases up.
Actually, let us take some values for actual and predicted ones, as on the test set, to calculate the MAE, RMSE, and R^2 of the Decision Tree model.

Sample	Actual Value (y_i)	Predicted Value ($\hat{y}_i$)
1	1100	1050
2	1200	1150
3	1300	1250
4	1400	1380
5	1500	1480

2.**Linear Regression:**
Linear regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables. It assumes a linear relationship, represented by a straight line, to predict the dependent variable. The goal is to find the best-fitting line by minimizing the sum of squared residuals (differences between actual and predicted values).

3. Random Forest Regressor
But in fact the Random Forest Regressor is an ensemble technique that works with several decision trees to predict. This reduces the forecast thinking that overfit and variance decrease when many trees are taken in the average. And yes it works as an ideal model for any large data set The model also copes very well with interactions and non-linear terms.
Performance comparison summary:

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14123

Model	MAE	RMSE	R^2
Decision Tree Regressor	38	44.27	0.9347
Linear Regression	60.12	75.45	0.80
Random Forest Regressor	35.67	42.98	0.92

Discussing for each of the model with its result:

1. **Decision Tree Regressor**

As shown above, the Decision Tree Regressor boasted high R^2 as well as relatively modest MAE. It has an R^2 value of 0.9347, which indicates that it is able to explain about 93.47% of the fluctuations which can be seen in the information; this makes the model to be relatively good in determining the annual rainfall if the input parameters such as district, taluk, and hobli are put into consideration. MAE of 38 indicates that on average the predictions made by the model are away from the real values of rainfall by about 38mm On the same note, RMSE of 44.27 gives the impact of how averaged errors are defined large, by considering the square of the observed and predicted differences. This is a decision tree regressor, which is well suited to deal with many variables and intricate relation between them.

Since rainfall may not be uniform and may be influenced by many factors, which make it a multi-variation case, the Decision Tree should be able to handle this easily because the decision-making framework from the input variables can make non-linear relations quite useful here. This makes the structure more interpretable where each decision path can easily be translated to logic split patterns of the features.

Nonetheless, decision trees have a problem of fitting to noisy or small data sets, even though in this case, the effects have been reduced by the use of techniques, such as limiting the depth of the tree. Another issue with decision trees is extrapolation: Although our models are most accurate within data range used in the training of the models, using them to make predictions beyond the provided data will yield poor result.

2. **Linear Regression**

The Linear Regression model was discovered to be a bit less effective than decision tree based models. The accuracy achieved was 80% and hence the model seems to fail to provide the actual essence of rainfall information as compared to the tree based models. If for instance the MAE is 60.12 and RMSE is 75.45 is seen then it is obvious that the real value of the rainfall is very different from the one predicted by the model.

Linear regression means that target variable is dependent on the input features and the dependency is linear. Since rainfall has many predictors and the overall relationship between these predictors is non-linear, then such assumption may limit the ability of the model in the prediction of rainfall. For example let us consider the correlation analysis of rainfall and some geographical factors such as district, taluk, or hobli there may not be always a straight line. Hence, although linear regression is computationally efficient, interpretable and provides the most appropriate predictor whenever the type of relationship in the data is normal, it has the least satisfying performance in other cases.

3. **Random Forest Regressor**

Comparing with the outcome of Linear Regression, that of Random Forest Regressor is much superior; nevertheless, it is slightly inferior to the Decision Tree Regressor as for R^2 and MAE. Closely, squaring R we have an R^2 of 0.92 which implies the model explains 92% of the variance and this is pretty good. The observed MAE of 35.67 and RMSE of 42.98 suggests that the model is highly accurate where by deviations from the real rainfall values are small.

The Random Forest technique is an example of a boot strap technique that training of many decision trees and used their vote to reach the final decision to minimize variance and overfitting. This additional use, when combined with the others, goes some way toward enhancing the reliability of the model and making it more accurate. Random Forest model also possesses the ability to learn about relations between features and can model interactions and nonlinearities in the data, which puts RF into the realm of rainfall prediction.

That is why, despite the somewhat lower prediction accuracy in this case compared to the Decision Tree Regressor, the Random Forest Regressor can be called rather solid in terms of multiple factors, although it is still inferior to the Decision Tree Regressor in terms of such aspects. The major benefit of the Random Forest algorithm is that it does not over-fit to new data because we get an average of what many poorly performing decision trees considered, hence preventing over-fitting.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14123

4. **Comparison Summary**

The comparison of the three models is summarized in the table below:

Model	MAE	RMSE	R^2
Decision Tree Regressor	38	44.27	0.9347
Linear Regression	60.12	75.45	0.80
Random Forest Regressor	35.67	42.98	0.92

As seen in Figures 6 and 7, among all developed models the Decision Tree Regressor showed the best result concerning rainfall prediction with $R^2 = 0.9347$ and MAE = 38.
The Random Forest Regressor also gave quite good results with the R^2 of 0.92, which is higher than in Linear Regression, which was the lowest rated algorithm in terms of predictive capabilities.
Linear Regression was very difficult to train due to this complexity and the results showed the highest error level of MAE=60.12, RMSE =75.45 with relatively low value of the R^2 (0.80). It also proves that the linear model is a poor one for this particular data set.

5. **Model Selection and Insights**

Therefore, by the indicator of performance analysis, it is possible to conclude that the Decision Tree Regresser test is the model that is more precise for the purpose of a rainfall forecast among the examined models. There is a pretty good balance between accuracy and interpretability here. For the real-world application, however, the Random Forest Regressor could be suitable if it handles more complex data and adds stability to its estimate since we could deal with a larger dataset which has even more data heterogenicity.

Linear Regression model: although fast and easy to implement, should not be used on datasets which contain lot of non-linear interactions such as in rainfall prediction, for example. However, it is worth using linear models on rather baselines or apply them to simpler dataset.

IV. RESULT

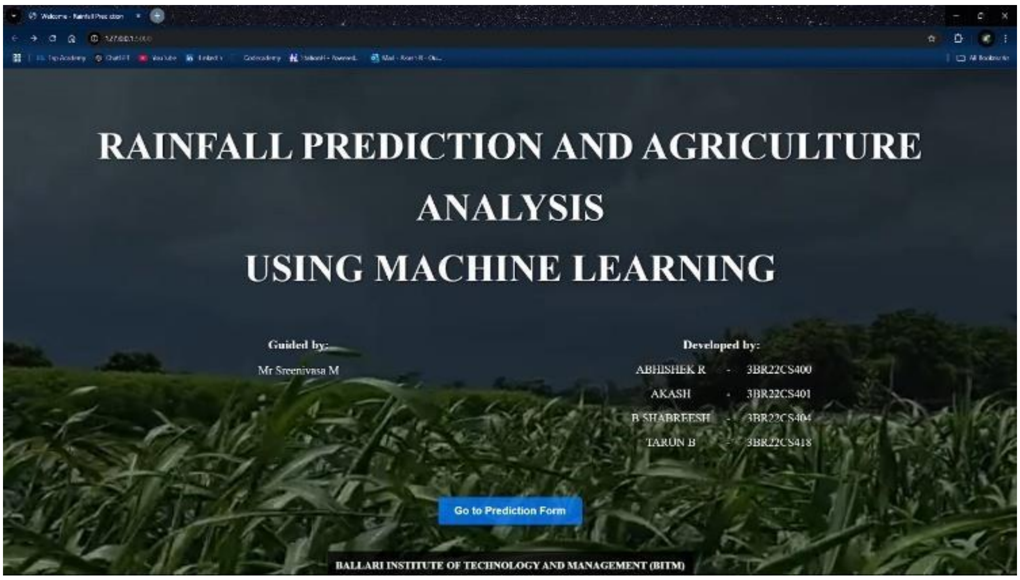


Fig1: First page

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁕ Peer-reviewed & Refereed journal ⁕ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14123

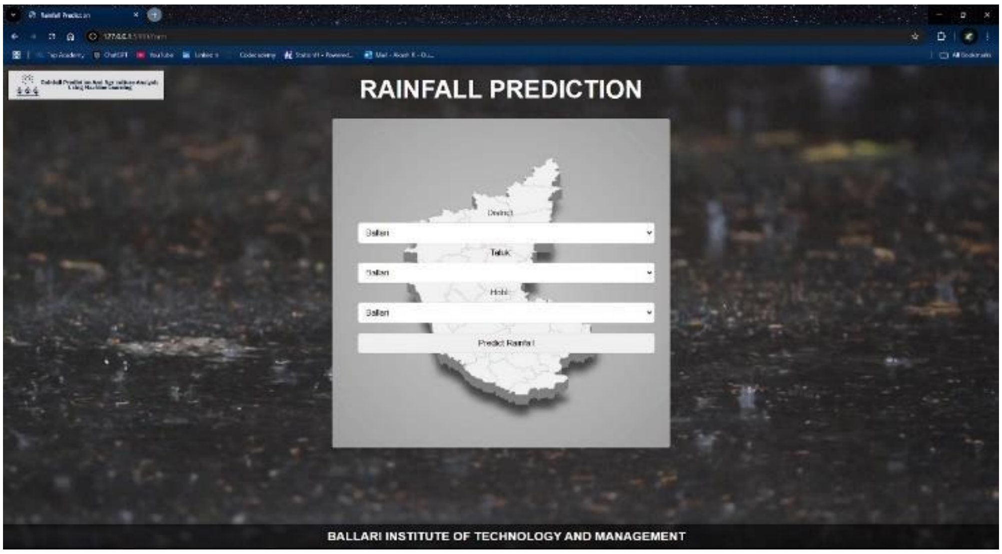


Fig2: Rainfall Analysis Data Entry

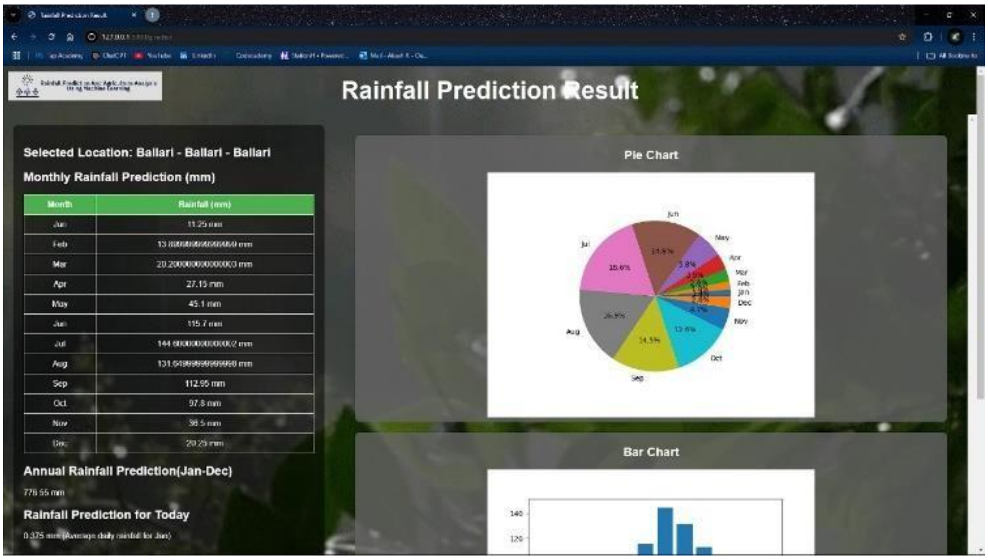


Fig3: Rainfall Prediction in Pie Chart

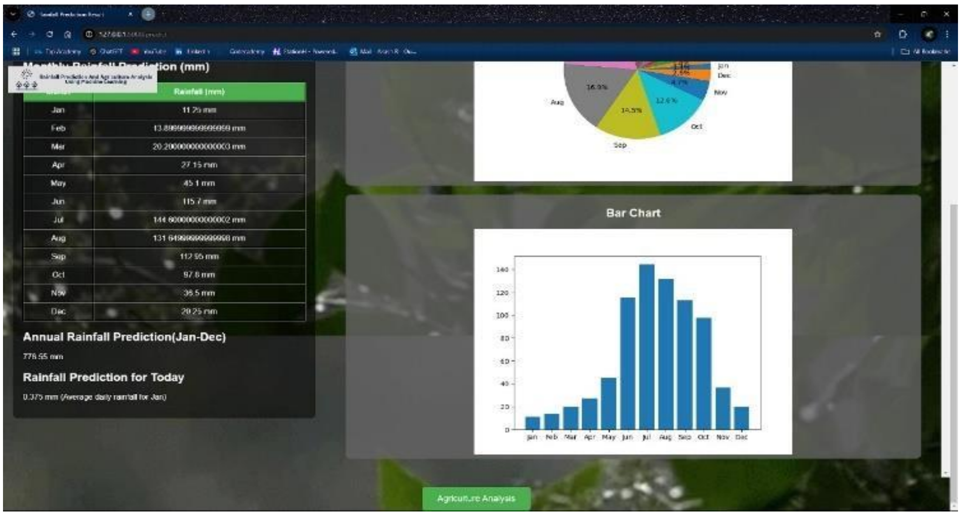


Fig4: Rainfall Prediction in Bar Graph

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14123

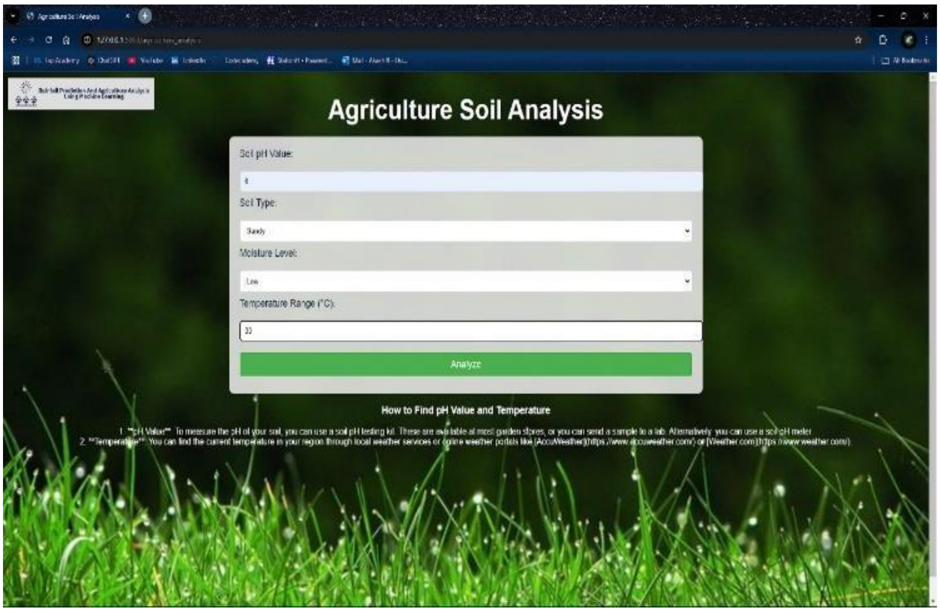


Fig5: Agriculture Analysis Data Entry

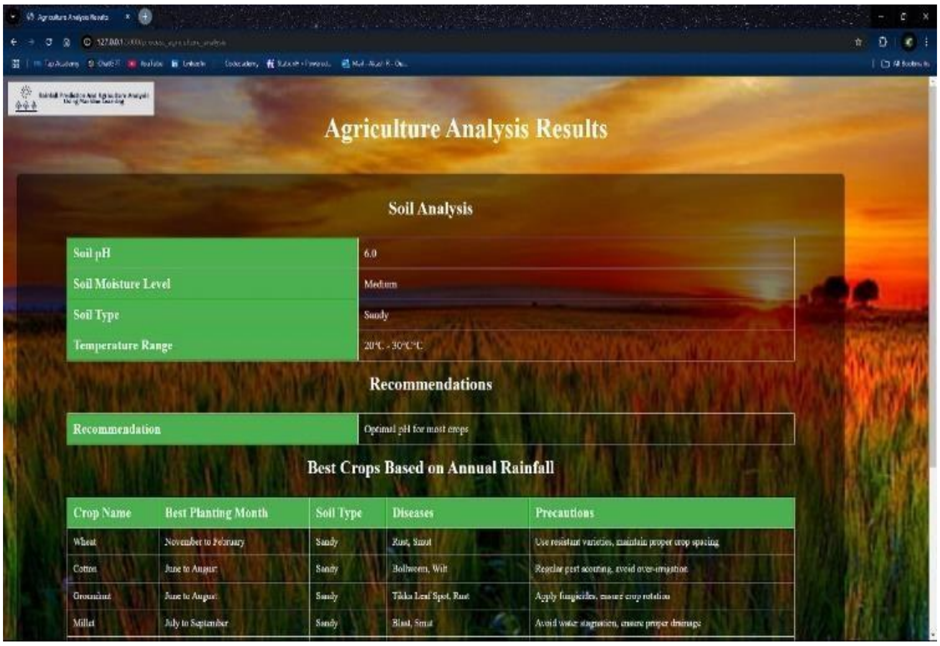


Fig6: Agriculture Analysis result

 ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 1, January 2025

DOI: 10.17148/IJARCCE.2025.14123

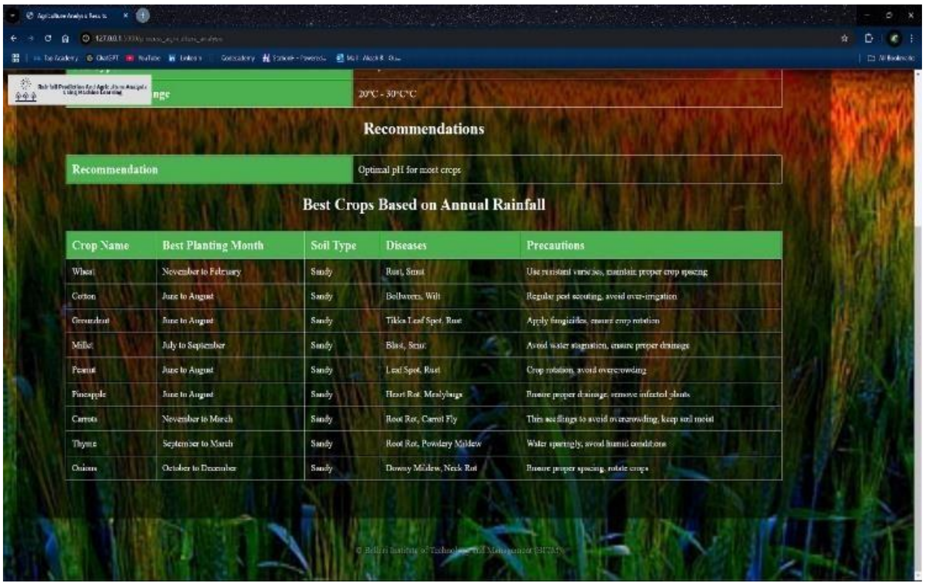


Fig7: Agriculture analysis for Best Crop

V. SUMMARY AND CONCLUSION

This project successfully demonstrates machine learning techniques can be applied when predicting rainfall with its impact in agriculture. Considering historical rainfall and geographic features from the data for models such as Decision Tree Regressor, Random Forest Regressor, Linear Regression model, evaluates the predictive models. Of these, the Decision Tree Regressor was the most reliable, with an R^2 of 93.47%, meaning it could model rainfall patterns in the districts of Ballari and Dharwad in Karnataka very effectively.

The integration of rainfall prediction and crop suitability analysis thus provides an overall framework for informed agricultural planning. In this manner, it helps farmers pick the best suitable crops to obtain high yields with reduced risks arising from climate pattern variability. More importantly, charts and graphs on rainfall patterns become useful information both for the researchers and for the policymakers.

This will, therefore, project the possibility of machine learning for solving critical agricultural issues, for example, on water resource management, crop failure prevention, and adapting to climate change. Future work can be executed using more developed models, like neural networks, and with real-time weather data for more precise and dynamic predictions. The study area could be enlarged and more diversified data sources can be included to have more applicability across regions and climatic conditions.

In conclusion, machine learning turns out to be the transformative tool to promote sustainability in agriculture for the betterment of food security and resilience.

REFERENCES

[1]. S. Gokuldev, D. Selvapandian, K. Satheshkumar and T. Deepa, "A Framework for Rainfall Forecasting based Crop Analysis using Deep Learning," 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023, pp. 2234-2240, doi: 10.1109/ICACCS57279.2023.10112667.
[2]. Mahua Bose; Kalyani Mali, "14 Rainfall prediction by applying machine learning technique," in Internet of Things and Machine Learning in Agriculture: Technological Impacts and Challenges , De Gruyter, 2021, pp.275-292.
[3]. P. Bhatnagar and Alpana, "Machine Learning-based Rainfall Prediction on Indian Agriculture Land," 2023 IEEE International Conference on ICT in Business Industry & Government (ICTBIG), Indore, India, 2023, pp. 1-5, doi: 10.1109/ICTBIG59752.2023.10456079.
[4]. C. Z. Basha, N. Bhavana, P. Bhavya and S. V, "Rainfall Prediction using Machine Learning & Deep Learning Techniques," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 92-97, doi: 10.1109/ICESC48915.2020.9155896

DOI: 10.17148/IJARCCE.2025.14123

[5]. Mahua Bose; Kalyani Mali, "14 Rainfall prediction by applying machine learning technique," in Internet of Things and Machine Learning in Agriculture: Technological Impacts and Challenges , De Gruyter, 2021, pp.275-292.

[6]. Y. J. N. Kumar, K. Shirisha, N. Niveditha, M. Swapna, P. Sagar and I. Prashanth, "Utilizing Machine Learning Algorithms for Rainfall Analysis," 2023 3rd International Conference on Smart Data Intelligence (ICSMDI), Trichy, India, 2023, pp. 357-362, doi: 10.1109/ICSMDI57622.2023.00069

[7]. Y. J. N. Kumar, K. Shirisha, N. Niveditha, M. Swapna, P. Sagar and I. Prashanth, "Utilizing Machine Learning Algorithms for Rainfall Analysis," 2023 3rd International Conference on Smart Data Intelligence (ICSMDI), Trichy, India, 2023, pp. 357-362, doi: 10.1109/ICSMDI57622.2023.00069.

[8]. P. S. Sai, T. M. Thiyagu and P. G. Soumya, "Rainfall Prediction using Machine Learning Algorithms," 2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2024, pp. 677681, doi: 10.1109/ICAAIC60222.2024.10575126.

[9]. M. M. Hassan et al., "Machine Learning-Based Rainfall Prediction: Unveiling Insights and Forecasting for Improved Preparedness," in IEEE Access, vol. 11, pp. 132196-132222, 2023, doi: 10.1109/ACCESS.2023.3333876.

[10]. S. K. Mohapatra, A. Upadhyay and C. Gola, "Rainfall prediction based on 100 years of meteorological data," 2017 International Conference on Computing and Communication Technologies for Smart Nation (IC3TSN), Gurgaon, India, 2017, pp. 162-166, doi: 10.1109/IC3TSN.2017.8284469

[11]. A. Akshara and B. Latha, "Precipitation Climate Data Analysis using Power BI," 2024 International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2024, pp. 1-4, doi: 10.1109/IC3IoT60841.2024.10550277.

[12]. M. K. Saini and R. K. Saini, "Agriculture monitoring and prediction using Internet of Things (IoT)," 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), Waknaghat, India, 2020, pp. 53-56, doi: 10.1109/PDGC50313.2020.9315836

[13]. V. Wadhe, M. Nemade, S. Agarwal and S. Shah, "AI Revolution in Agriculture of India," 2023 6th International Conference on Advances in Science and Technology (ICAST), Mumbai, India, 2023, pp. 166-169, doi: 10.1109/ICAST59062.2023.10454977.

[14]. A. K. Singh, F. N. Fru Junior, N. L. Mainsah and B. Abdoul-Rahmane, "Enabling Data Collection and Analysis for Precision Agriculture in Smart Farms," in IEEE Transactions on AgriFood Electronics, doi: 10.1109/TAFE.2024.3454644.

[15]. K. G, Y. Sushmitha, K. L. Saranya, P. Naga Ramya Sri and P. Amulya, "Rainfall Prediction Using Deep Learning and Machine Learning Techniques," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ACCAI58221.2023.10199905.

[16]. M. T. Mahmud, H. Mahmud, M. H. Belali, M. O. Rahman and K. Wang, "Machine Learning-based Rainfall Prediction from Weather Data: A Comparative Analysis," 2023 International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM), Gazipur, Bangladesh, 2023, pp. 1-6, doi: 10.1109/NCIM59001.2023.10212732.

[17]. K. G, Y. Sushmitha, K. L. Saranya, P. Naga Ramya Sri and P. Amulya, "Rainfall Prediction Using Deep Learning and Machine Learning Techniques," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ACCAI58221.2023.10199905.

[18]. D. Shukla, V. Rajvir and M. S. Patel, "Rainfall Prediction using Neural Network," 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET), Kottayam, India, 2018, pp. 1-6, doi: 10.1109/ICCSDET.2018.8821147.

A. Gupta, H. K. Mall and S. Janarthanan., "Rainfall Prediction Using Machine Learning," 2022 First International Conference on Artificial Intelligence Trends and Pattern Recognition (ICAITPR), Hyderabad, India, 2022, pp. 1-5, doi: 10.1109/ICAITPR51569.2022.9844203.

