Database & Analytics Programming (MSCDAD_C)
Continuous Assessment - 2

Project Report

"Weather Based Crop Recommendation in India"

Anakha Sreekumar
X22237691

Sushma Suresh Hebbalakar
x22224122

Master's in Data Analytics
National College of Ireland
17-12-2023

Weather Based Crop Recommendation in India

Sushma Suresh Hebbalakar
M.Sc in Data Analytics
National College of Ireland
x22224122

Anakha Sreekumar
M.Sc in Data Analytics
National College of Ireland
x22237691

***Abstract*—The field of research in agriculture is expanding day by day. Crop prediction in particular is very important in agriculture and primarily depends on the soil and environmental factors, such as temperature, humidity, rainfall and soil contents. The farmers were the one who decided the crop to be cultivated and monitored its growth in the past. But now a days it is challenging for them to predict these thing due to the rapid change in the climate and environmental conditions. This has led to the replacement of prediction tasks with the machine learning techniques in the recent years, some of which have been used in this work to predict crops to be cultivated.**

I. Introduction

Farming is backbone of most developing nations, and farmers play a vital role in providing food security to their respective countries. However, farmers face numerous challenges, such as unpredictable weather patterns, soil degradation, pests, and diseases, which can significantly impact crop production. To overcome these challenges, farmers need to have access to relevant information and data-driven insights to make informed decisions about their farming practices.

Our goal is to find the trend in the temperature and rainfall studying the collected datsets over 116 years and further analyse features that influence that better suits the crop growth and recommend crops. One of the sources for our datasets is Indian government site. We have collected the rainfall and temperature data over the period of 1901 and 2016 from gov.in and the crop recommendation data from the WHO website.

In the recent time there is a drastic change in the temperature, and it is affecting the agriculture field. As a result, it becomes difficult for the farmers to predict which crop is suitable for the season.The use of machine learning-based Agricultural Crop Recommendation System has the potential to bring about a substantial increase in agricultural productivity.

By providing farmers with data-driven recommendations, this system can assist them in making informed decisions that are conducive to better yield, thus leading to more efficient and profitable farming practices, reduce costs, and increase efficiency in agricultural activities. It can help farmers to optimize their crop production, minimize environmental impacts, and ensure sustainable farming practices.

This system can also help to bridge the gap between farmers and technology, allowing farmers to access advanced technological tools and expertise that were previously un available. In this paper, we will discuss the implementation of the Agricultural Crop Recommendation System using machine learning, its benefits, and potential impact on the agricultural industry

II. Related Work

A lot of research has been done on suggesting which crops are best for certain types of soil and when would be right time to grow based on temperature and rainfall. Many studies highlight how knowing about soil and its effects on crops is crucial for making smart decisions about what to plant. In this section, We'll give you a quick overview of the main studies and methods in this area According to [2] this paper recommends different crops of India and their analysis is given using machine learning algorithms. The Analysis has been performed on six types of machine learning algorithms and out of these six algorithms XGBoost achieved best accuracy result.Such performance metric will definitely help farmers in making decision that what are values required of each first seven attributes viz. N, P, K, temperature, humidity, ph, rainfall in order to grow any particular crop. This in turn will indirectly help in boosting economy of the country [2].

Paper [3] explains, crop recommendation system is an important tool for farmers that can help them make data-driven decisions on which crops to grow based on various factors such as soil type, climate, and historical crop yield data.And the methodology for building a crop recommendation system involves several steps, including data collection, pre-processing and cleaning, feature selection and engineering, model selection and training, model evaluation and optimization, and deployment. By following this methodology, developers can build accurate and reliable crop recommendation systems that can provide farmers with personalized recommendations to optimize their crop yields and improve their overall farming practices [3]

. The paper [5], is about Machine learning applied in the proposed system to analyze crop yield under various weather situations. Datasets on crop growth have been gathered for training and testing purposes. They found that using a variety of outputs, the RF algorithm is a powerful technique for analyzing crop productivity under diverse climatic situations. [5]

The paper [1] , the system's recommendations include crop recommendations and crop sustainability predictions. By giving information that normal farmers do not monitor, the suggested approach helps farmers choose the optimal crop by reducing the likelihood of crop failure and production. The

system employed the Python programming language and two algorithms: linear regression and neural networks [1].
In Paper [6] they have explained the necessary parameter data, including temperature, humidity, pH, and the ability to detect rain, which allows us to calculate the amount of rainfall. For the purposes of subsequent crop prediction, these values are used [6].
This Paper [7] proposes a yield prediction system for farmers in India, which includes a mobile application that uses machine learning algorithms to suggest profitable crops or predict crop yields based on user inputs such as location and soil type,System not only offers reliable outcomes but also recommends the most favorable time to apply fertilizers to boost crop yield. By leveraging this information, farmers can ensure that they use fertilizers at the appropriate time, leading to more sustainable and productive farming practices [7].
[4] The authors propose an Internet of Things and AIML system for soil testing using sensors to maintain crop health and lower the probability of soil degradation. They suggest using various sensors such as soil temperature, soil moisture, pH, and N, P, K to monitor soil parameters. The data collected by the sensors deployed in the agricultural fields is analyzed through machine learning algorithms like random forest. The insights generated through this analysis enable the system to recommend the best crops that are suitable for growth under the given conditions, thereby providing farmers with data-driven insights to make informed decisions. [4]

III. Methodology

Crop recommendation system is designed to assist farmers in making data-driven decisions about which crops to grow based on several factors such as soil type, climate, and historical crop yield data. The methodology for creating a crop recommendation system involves the following steps:

A. Dataset description / Data Collection:

The initial step is to take data on different factors that affect crop growth and yield, such as soil type, climate, weather patterns, topography, and historical crop yield data.
1) Dataset 1 – Maximum temperature data from https://data.gov.in/ Raw data was collected in xml formats from the Indian Government open API.
2) Dataset 2 - Minimum temperature data from https://data.gov.in/ Raw data was collected in xml formats from the Indian Government open API.
3) Dataset 3 - Rainfall data from https://data.gov.in/ Raw data was collected in xml formats from the Indian Government open API.
4) Dataset 4- Crop recommendation data from https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset
The Datasets were loaded to jupyter notebook import functions written to read xml and csv files and then were loaded to postgres using the packages and libraries available in python to write to and read from sql .
Dataset 1 consists of Year wise , Month wise , Annual and Seasonal maximum temperature (C) from 1901-2016.
Dataset 2 consists of Year wise , Month wise , Annual and Seasonal minimum temperature in (C) from 1901-2016.
Dataset 3 consists of Year wise , Month wise , Annual and Seasonal rainfall in (mm) from 1901-2014.

Since we had all the real-time data and almost 116 ears data we thought this daa would be right to analyse the temperature and rainfall trend.All these datasets are used to analyse and predict future max and min temperatures along with predicted rainfall.
Crop recommendation dataset was also used to analyze and recommend the crop based on various factors which also includes temperature and rainfall as one of its factors.

B. Detailed Description of Data Processing:

Fig 1 shows the visual overview of data gathering and processing. We had the raw data in the form xml and csv using python we have done the ETL steps.As part of **Extraction**, we have extracted the data from xml and csv successfully using the python libraries available.For **Transformation** we have modified the data as per our requirements which will be helpful for doing Analysis and training the data.And as part of **Loading** we have loaded the transformed or processed data to Postgres Database.
Here for the project we have used 4 datasets which will help us in analysis the maximum temperature, minimum temperature, rainfall by analysis the data for over 116 years. We have used the Indian Government open API https://data.gov.in/ to gather these 3 datasets and https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset to gather crop recommendation dataset which has 2200 instances in total for 22 different crops.

C. Justifications of the implemented data processing :

As part of data processing we have transformed the data into our need. Below Table-1 shows the data we had before processing and Table-2 shows the data we had after processing. Since we had the data in the form which was not helpful for analysis we have transformed as shown. There were no null values present and we have not removed outliers since the data we collected were real time data and we wanted to train and analyse the same. And for crop recommendation dataset based on the models we are using we have applied Standard scalar and Min max scalar for better performance of the model.

	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	JAN-FEB	MAR-MAY	JUN-SEP	OCT-DEC
0	1901	22.40	24.14	29.07	31.91	33.41	33.18	31.21	30.39	30.47	29.97	27.31	24.49	28.96	23.27	31.46	31.27	27.25
1	1902	24.93	26.58	29.77	31.78	33.73	32.91	30.92	30.73	29.80	29.12	26.31	24.04	29.22	25.75	31.76	31.09	26.49
2	1903	23.44	25.03	27.83	31.39	32.91	33.00	31.34	29.98	29.85	29.04	26.08	23.65	28.47	24.24	30.71	30.92	26.26
3	1904	22.50	24.73	28.21	32.02	32.64	32.07	30.36	30.09	30.04	29.20	26.36	23.63	28.49	23.62	30.95	30.66	26.40
4	1905	22.00	22.83	26.68	30.01	33.32	33.25	31.44	30.68	30.12	30.67	27.52	23.82	28.30	22.25	30.00	31.33	26.57

Table 1: Before processing dataset

	YEAR	MONTH	max_temp
0	1901	JAN	22.40
1	1902	JAN	24.93
2	1903	JAN	23.44
3	1904	JAN	22.50
4	1905	JAN	22.00

Table 2: After Processing datset

D. Justifications for the choice of technologies used: Justifications for the choice of technologies, such as the programming language Python and the use of PostgreSQL databases and some libraries are as below:

1.**Programming Language - Python**: Python was chosen as the programming language due to its versatility, readability, and extensive ecosystem of libraries. It is well-suited for data analysis, machine learning, and scripting tasks.

2.**XML Parsing - 'xml.etree.ElementTree':** The 'xml.etree.ElementTree' module is a standard library in Python for working with XML data. It provides a simple and efficient way to parse and manipulate XML structures, making it suitable for handling XML-based data.

3.**Data Manipulation and Analysis - 'pandas':** The 'pandas' library is a powerful tool for data manipulation and analysis. Its DataFrame structure allows for efficient handling and processing of tabular data, which is often essential in data science projects.

4.**Data Visualization - 'matplotlib' and 'seaborn':** 'matplotlib' and 'seaborn' are popular data visualization libraries in Python.They provide a wide range of customizable plots and charts, making it easy to create visual representations of data patterns and insights.

5.**Database Interaction - 'sqlalchemy':** Reasoning: 'sqlalchemy' is a widely used SQL toolkit and Object-Relational Mapping (ORM) library for Python. It allows for easy integration with different databases, including PostgreSQL, providing a flexible and efficient way to interact with databases in a Pythonic manner.

6.**Machine Learning - 'scikit-learn':** 'scikit-learn' is a comprehensive machine learning library in Python. It includes a variety of tools for regression, classification, clustering, and more. The use of 'LinearRegression', 'DecisionTreeRegressor', and 'RandomForestRegressor' suggests a focus on predictive modeling.

7.**Model Evaluation - 'sklearn.metrics':** The inclusion of metrics from 'sklearn.metrics', such as mean squared error, mean absolute error, and R-squared score, indicates a commitment to evaluating and measuring the performance of machine learning models.

8.**Math Operations - 'numpy':** 'numpy' is a fundamental library for numerical operations in Python. Its array-oriented computing capabilities are valuable for mathematical operations, and it is commonly used in conjunction with other libraries in the data science and machine learning ecosystem.

9.**Data Splitting - 'train_test_split':** The use of 'train_test_split' from 'sklearn.model_selection' suggests a commitment to best practices in machine learning, including the separation of data into training and testing sets to evaluate model performance effectively.

10.**Database : Postgres :** PostgreSQL is a powerful and reliable open-source database management system. It's known for its stability, support for complex queries and large datasets, and adherence to ACID principles for data integrity. PostgreSQL is highly customizable, scalable, and supports SQL standards. With a strong community, security features, and compatibility with various data types, including advanced features like JSON support and spatial data handling, PostgreSQL is a versatile choice for projects with diverse data management needs.

E. Diagrams providing a visual overview:
Below Figure 1 shows the visual overview of the data gathering and processing.Figure 2 shows the analysis flow.

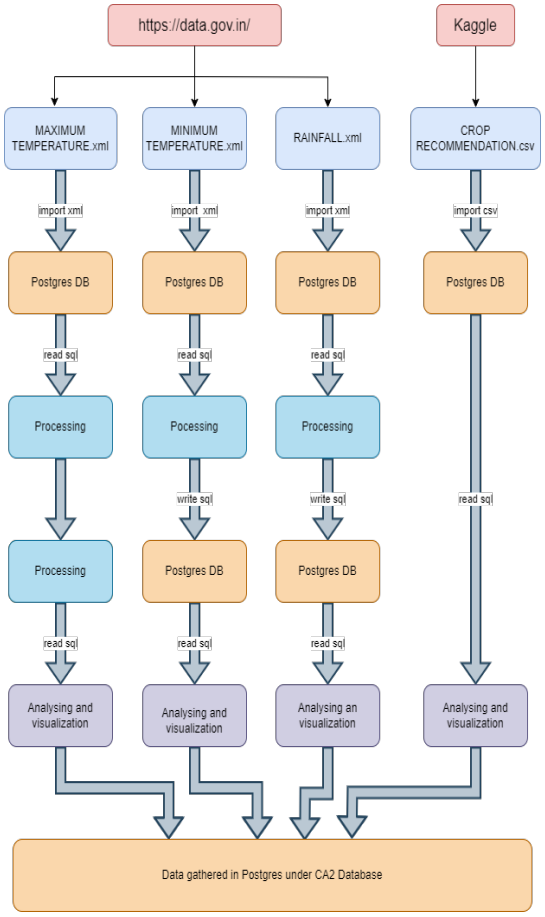

Fig 1: Diagram for data gathering and processing

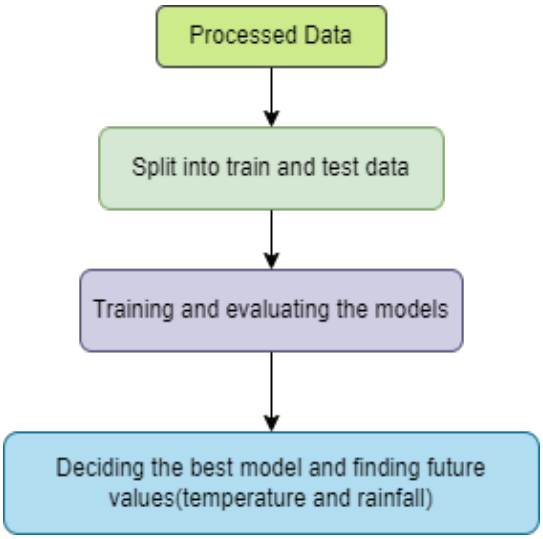

Fig 2: Diagram for data analysis

IV. Results and Evaluation

A. Results

The result and evaluation of the final process is to load the processed/tranformed data into PostgreSQL and from there data has been fetched and stored in the dataframe of the python where the whole visualizations on the data has been performed. And finally data is interpreted and analyzed from the visualization. Visualization has been carried out on each of the dataset separately based on which interpretation and analysis has been performed.

1.Visualization on maximum temperature over years and months

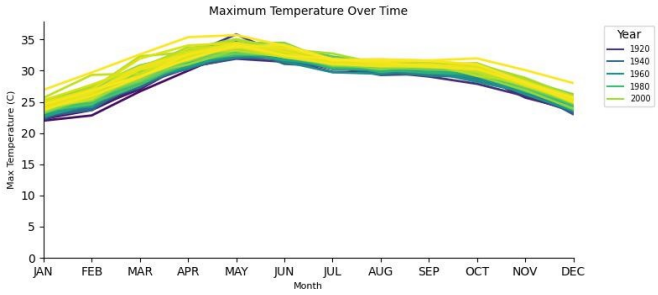

Figure 3: Maximum temperature by year and month

We can see that the maximum temperature range between 22° to 35° and during period of 20 years from 2000 there was highest temperature. I have used lineplot to show how the maximum temperature was from year 1901 to 2016 every month

2.Visualization on minimum temperature over years and months

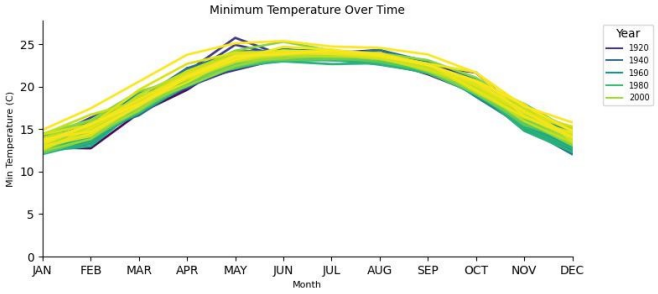

Figure 4: Minimum temperature by year and month

We can see that the minimum temperature range between 12° to 25° and during period of 20 years from 1920 there was lowest temperature.I have used lineplot to show how the minimum temperature was from year 1901 to 2016 every month

3.Visualization on rainfall over years and months

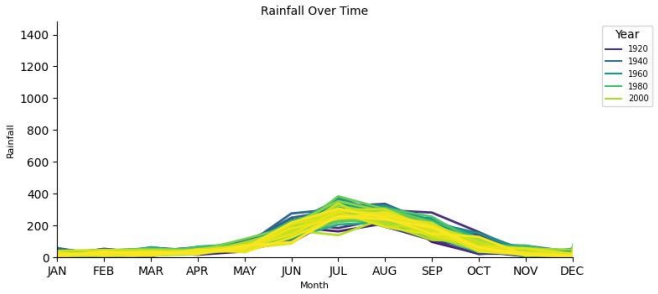

Figure 5: Rainfall (in mm) by year and month

We see that the rainfall range between 0 mm to 400 mm and during period of 20 years from 2000 there was highest rainfall.I have used lineplot to show how the raifall was from year 1901 to 2014 every month

4.Visualization on maximum temperature over years seasonally

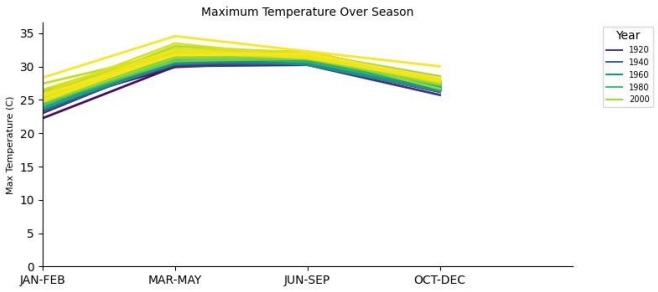

Figure 6: Maximum Temperature (C) by year and season

It is evident that from March to May we have maximum temperature which range between28° to 35° and this period will be considered as Summer season in India as we have

high temperature over these months.

5.Visualization on minimum temperature over years seasonally

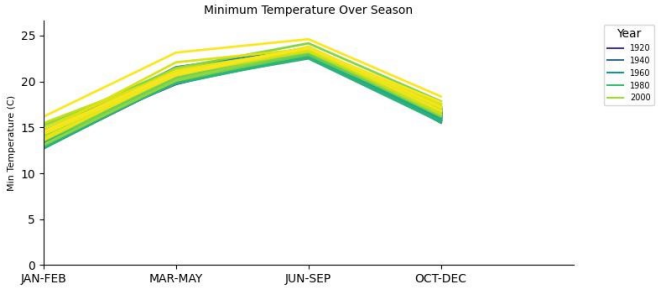

Figure 7: Minimum Temperature (C) by year and season

We see that from Jan to Feb we have minimum temperature which range between12° to 16° and this period will be considered as Winter season in India as we have low temperature over these months we also consider December month as part of winter season but as per the data we are seeing low temperature duing Jan and February considering all 116 year data.

6.Visualization on comparing maximum temperature over years and the average temperature over years

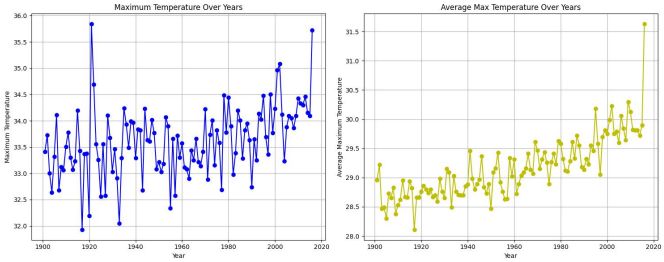

Figure 8: Maximum Temperature(C) v/s Average maximum temperature by year

Figure 8 shows that we had maximum temperature of equivalent to 35.9 in 1921 and second highest during 2016 around 35.6 but on an average the maximum temperature was seen in 2016. The Figures also says that the average maximum temperature is increasing over years from 1901 to 2016.

7.Visualization on comparing minimum temperature over years and the average temperature over years

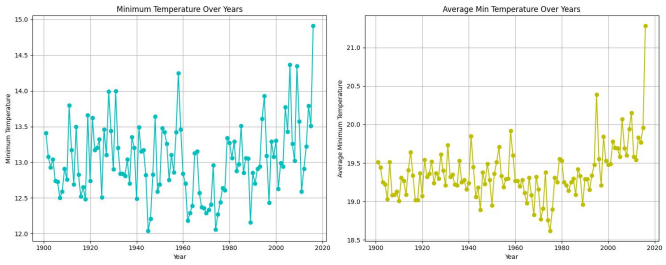

Figure 9: Minimum Temperature(C) v/s Average minimum temperature by year

Figure 9 shows that we had lowest temperature of equivalent to 11.9 in 1945 and second highest during 1974 around 11.6 but on an average the minimum temperature was seen in 1975. The Figures also says that the average minimum temperature is increasing over years from 1901 to 2016.

8.Visualization on average rainfall

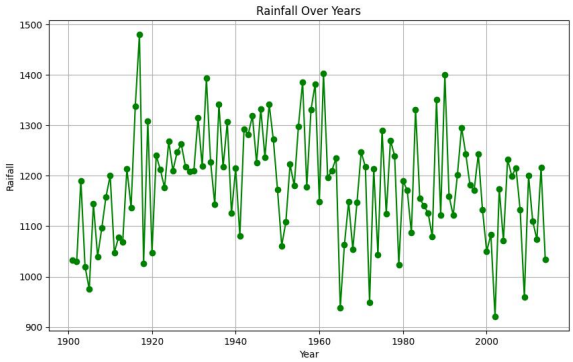

Figure 10: Average rainfall by year

Figure 10 shows that we had highest rainfall in 1917 and lowest rainfall in 2002. The Figures also depicts that the average rainfall is decreasing over years from 1901 to 2014.

9.Visualization on maximum and minimum temperature required for 22 different kind of crops

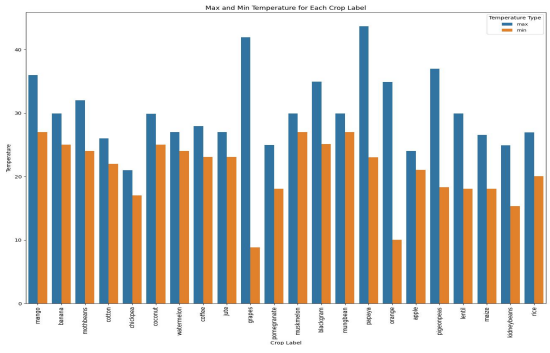

Figure 11: Maximum and Minimum temperature required for crops

Figure 11 shows the maximum and minimum temperature required for crops to grow based on the dataset we collected.

It reveals that grapes can grow in both low and high temperature while other crops have different temperature range

10.Visualization on amount of nitrogen required in soil for 22 different kind of crops

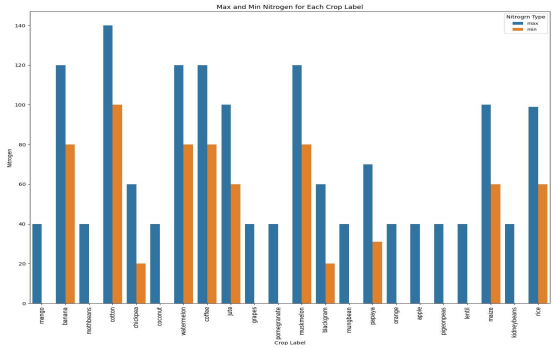

Figure 12: Maximum and Minimum Nitrogen content required for crops

Figure 12 visualises that the maximum and minimum Nitrogen required in soil for crops to grow based on the dataset we collected. It shows that crops ”Banana”, ”Cotton”, ”Chickpea”, ”Watermelon”, ”Coffee”, ”Jute”, ”Muskmelon”, ”Blackgram”, ”Papaya”, ”Maize” and ”Rice” require the some amount nitrogen to be present in soil while for others it’s not really required.

13.Visualization on amount of potassium required in soil for 22 different kind of crops

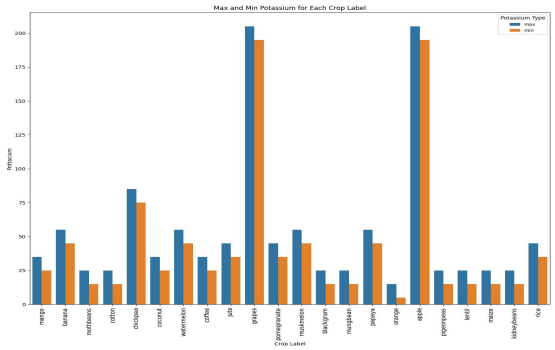

Figure 12: Maximum and Minimum Potassium content required for crops

Figure 12 visualises that the maximum and minimum Potassium required in soil for crops to grow based on the dataset we collected. It shows that all the crops require some amount of potassium to be present in the soil for them to grow well. Specially ”Apple” and ”Grapes” require more amount of potassium to be present in the soil compared to all others.

12.Visualization on amount of Phosphorus required in soil for 22 different kind of crops

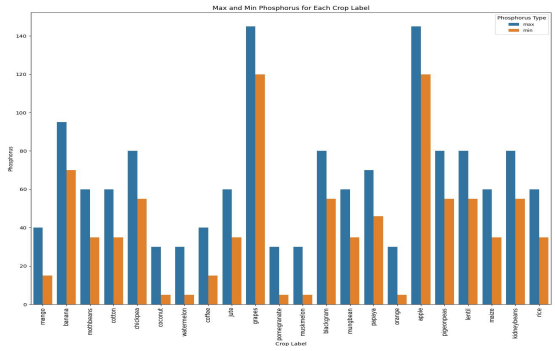

Figure 14: Maximum and Minimum Phosphorus content required for crops

Figure 14 shows that the maximum and minimum Phosphorus required in soil for crops to grow based on the dataset we collected. It shows that all the crops require some amount of phosphorus to be present in soil and the amount of Phosphorus required for ”Apple” and ”Grapes” is high.

B. Evaluation

1) Evaluation on maximum and minimum temperature: To predict the temperature and rainfall by year and month we decided to use LinearRegression ,DecisionTreeRegressor , RandomForestRegressor and tried to evaluate them by mean_squared_error, mean_absolute_error, r2_score and root_mean_squared_error. Linear regression is good for capturing linear relationships, while decision trees and random forests can handle more complex patterns and interactions. As we evaluated the models on all three we saw that random forests performed better giving accurate prediction on our datset. Below are the interpretations we drew from training the models.

Model/Interpretation	Random Forest Regression Metrics:	Decision Tree Regression Metrics:	Linear Regression Metrics:
Mean Squared Error (MSE):	0.584543494	0.797574713	9.741615508
Mean Absolute Error (MAE):	0.572877586	0.670574713	2.644284667
R-squared:	0.939543953	0.917511332	-0.00752052
Root Mean Squared Error:	0.764554441	0.893070385	3.121156117

Table 3: Interpretation on models to find maximum temperature

The provided metrices suggest that the Random Forest Regression model performs quite well on the evaluated dataset:

1.Mean Squared Error (MSE): This metric measures the average of the squared differences between the predicted and actual values. A lower MSE indicates better predictive performance.

2.Mean Absolute Error (MAE): MAE represents the average of the absolute differences between predicted and actual values. Like MSE, lower MAE values indicate better model

accuracy.

3.R-squared: R-squared quantifies the proportion of variability in the target variable that is explained by the model. A value of 0.9395 indicates a high goodness of fit, suggesting that the model captures a significant portion of the variance in the data.

4.Root Mean Squared Error (RMSE): RMSE is the square root of the MSE, providing a measure of the average magnitude of errors. A lower RMSE is indicative of better model performance.

In summary, the **Random Forest Regression model**, as indicated by these metrics, appears to be effective in making accurate predictions on the given dataset.

The Residual Plot and actual_vs_predicted plot for random forest model is as shown below:

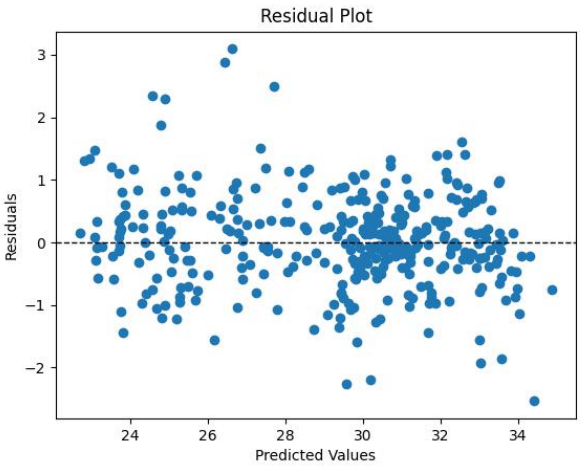

Figure 15: Residual plot for maximum temperature prediction using random forest model

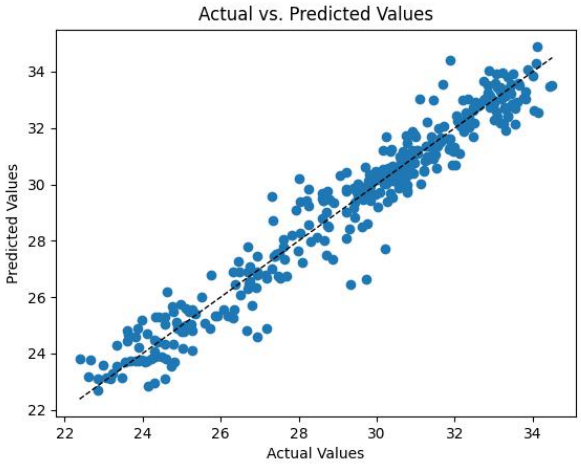

Figure 16: Actual v/s predicted values for maximum temperature using random forest model

Similarly to find the minimum temperature we have again evaluated all 3 models and found the below interpretations.

Model/Interpretation	Random Forest Regression Metrics:	Decision Tree Regression Metrics:	Linear Regression Metrics:
Mean Squared Error (MSE):	0.324300965	0.460199425	15.25945205
Mean Absolute Error (MAE):	0.431473276	0.515229885	3.485265076
R-squared:	0.978713586	0.969793505	-0.001597417
Root Mean Squared Error:	0.569474288	0.678380001	3.906334862

Table 4: Interpretation on models to find minimum temperature

In summary, the **Random Forest Regression model**, as indicated by these metrics, appears to be effective in making accurate predictions on the given dataset.

The Residual Plot and actual_vs_predicted plot for random forest model is as shown below:

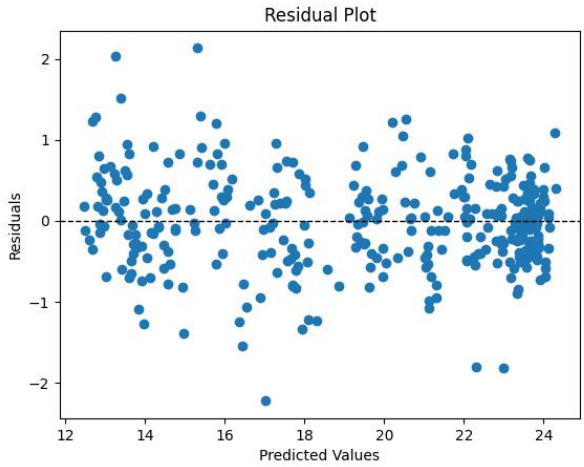

Figure 17: Residual plot for minimum temperature prediction using random forest model

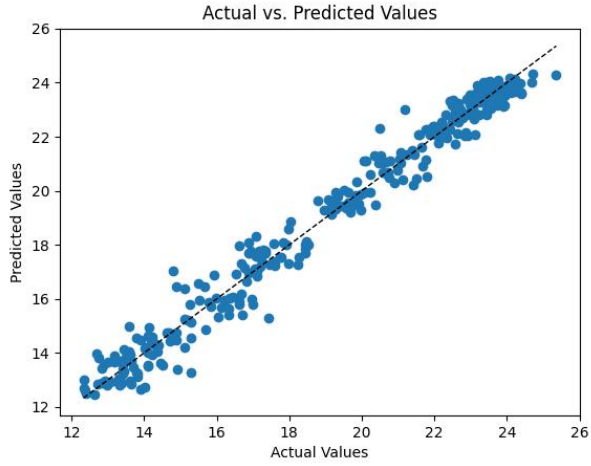

Figure 18: Actual v/s predicted values for minimum temperature using random forest model

Residual Plot: A residual plot helps us to check whether the residuals (the differences between observed and predicted values) exhibit a pattern. Ideally, residuals should be randomly

scattered around zero.

Randomness: Residuals should show no clear pattern, they should be evenly scattered. Constant Spread: The spread of residuals should be roughly constant across all predicted values.

Actual vs Predicted Plot: This plot helps you visually compare the model's predictions to the actual values. It provides insights into how well the model is capturing the underlying patterns in the data.

Linearity: The points should be close to a diagonal line, indicating a good fit. Spread: The spread of points around the line should be relatively uniform.

Our Figures 14 and 16 i.e., residuals plots are showing no patterns and are evenly scattered. and Figure 15 and 17 i.e., Actual_vs_predicted shows linearity and uniformity.

Below figures 18 and 19 shows the line plot drawn on future maximum and minimum temperature found using the random forest model we trained to predict the values.

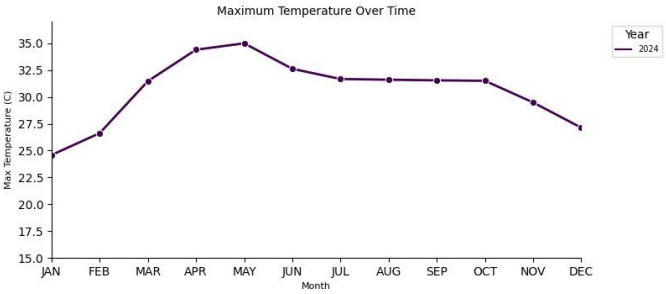

Figure 19: Maximum temperature predicted for future

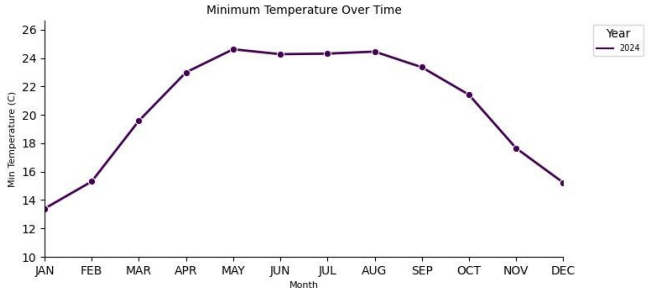

Figure 20: Minimum temperature predicted for future

2) Evaluation on crop recommendation dataset we have::
We have used 'Logistic Regression', 'TensorFlow', 'Decision Tree', 'Random Forest', 'Naive Bayes', 'SVM' models are being used to predict the crop to be grown, given, Nitrogen level in soil, Phosphorus level in soil, Potassium level in soil, ph, humidity, temperature and rainfall.
While training data in LogisticRegression we have done Standard scalar transformation on our dataset and then trained our model using scaled data so that the model performs better.
And before training SVM model we have applied MinMaxScaler() tranformation on training data and testing data and then trained and tested the model.
Below are the interpretations drawn after training the model.The Table 5 shows that both "SVM" and "Naive Bayes" have performed well and shown better results hence we can use these models to recommend the crop to be grown.Below Figure 21 shows the visual representation of accuracy of different models to recommend/predict the crop.

Models	Accuracy
'Logistic Regression'	96.90909091
'TensorFlow'	98.18181992
'Decision Tree'	85.81818182
'Random Forest'	99.45454545
'Naive Bayes'	99.63636364
'SVM'	99.63636364

Table 5: Interpretation of accuracy by different models to recommend the crop.

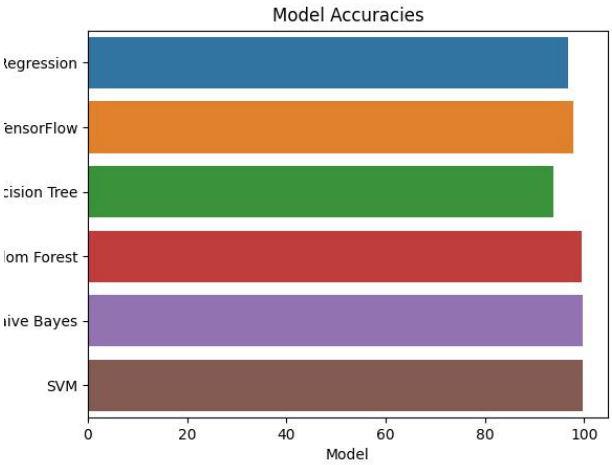

Figure 21: Different models accuracy to recommend crop.

V. CONCLUSION AND FUTURE WORK

A. Conclsion

The crop recommendation system developed in this project has showcased promising results in predicting suitable crops based on soil nutrients, climatic conditions, and environmental factors. By leveraging machine learning algorithms, we achieved a commendable accuracy in suggesting crops tailored to specific soil compositions and weather patterns. The model's ability to consider a multitude of features such as nitrogen, phosphorous, potassium levels, temperature, humidity, pH, and rainfall has significantly enhanced its predictive capabilities. This system serves as a valuable tool for farmers, aiding them in making informed decisions regarding crop selection, leading to improved yields and resource utilization.

Below Table 6, shows the insights we have drawn from analysing the temperature and shows the recommended crops based on the maximum temperature limit as all the crops can be grown considering the minimum temperature level we have seen. We are focusing on maximum temperature limit required to grow specific crop.

Among all crops ”Grapes”, ”Papaya”, ”Blackgram”, ”Pigeonpeas”, ”Mango”, ”Orange”, ”Lentil” can be cultivated anytime based on the analysis we drew from the dataset we had. And we see that March - Sep we have very limited crops that can be grown based on the datsets we have. Specially March- May has very few crops that can sustain during high temperature. And Oct - Feb is most suitable period to grow most of the crops.

With respect to Machine learning we have concluded to use Randomforest regression to predict the maximum and minimum temperature since it has performed better as shown in Table 4. And for crop recommedndation we have concluded to use either Naive Bayes or SVM as both them have shown better accuracy as seen in Table 5.

Crops	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec	All Season
Coconut	✓		✓	✓	
Grapes					✓
Pappaya					✓
Kidneybean	✓				
Blackgram					✓
Mungbean	✓		✓	✓	
Chickpea	✓				
Coffee	✓			✓	
Jute	✓			✓	
Muskmelon	✓		✓	✓	
Pigeonpeas					✓
Maize	✓				
Rice	✓				
Mothbeans	✓		✓	✓	
Mango					✓
Banana	✓		✓	✓	
Cotton	✓			✓	
Watermelon	✓			✓	
Pomegranate	✓				
Orange					✓
Apple	✓			✓	
Lentil					✓

Table 6: Overview on crops recommended based on their maximum temperature limit..

B. Future Work

As part of future work we can work on getting more appropriate dataset from real-time analysis using IOT and fetch the appropriate data which could help us in doing analysis. For weather prediction we can look at other factors that influence the weather. The wind, humidity and many more.

For crop recommendation we can look at other factors that will influence the better crop growth we can collect the data based on locations/regions to recommend better crops that suits the place. We can also work on finding the amount of pesticides we can apply for the crop based on the soil condition to get better results. There are more factors/features that are needed together for better growth of crops.

Overall we can work on gathering weather data based on location, Soil quality analysis, Mai components that effect the crop growth based on the crop we are growing for example., does the crop require more nitrogen or potassium or something else. If all these analysis is done in future we can build a more strong model that will help in better recommendation of crops.

REFERENCES

[1] Priyadharshini A, Swapneel Chakraborty, Aayush Kumar, and Omen Rajendra Pooniwala. Intelligent crop recommendation system using machine learning. In *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, pages 843–848, 2021.

[2] Ajay Agarwal, Sartaj Ahmad, and Adesh Pandey. Crop recommendation based on soil properties: A comprehensive analysis. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–6, 2023.

[3] D. Balakrishnan, Anumula Praneeth Kumar, Kristipati Sai Kiran Reddy, R. Ravindra Kumar, K. Aadith, and Sudarsi Madhan. Agricultural crop recommendation system. In *2023 3rd International Conference on Intelligent Technologies (CONIT)*, pages 1–5, 2023.

[4] Gaurav Chauhan and Alka Chaudhary. Crop recommendation system using machine learning algorithms. In *2021 10th International Conference on System Modeling Advancement in Research Trends (SMART)*, pages 109–112, 2021.

[5] V. Geetha, A. Punitha, M. Abarna, M. Akshaya, S. Illakiya, and A.P. Janani. An effective crop prediction using random forest algorithm. In *2020 International Conference on System, Computation, Automation and Networking (ICSCAN)*, pages 1–5, 2020.

[6] C. V. Neha Niharika, K. Mothish Kumar, K. Harika, V. Venkatesh, and A. Yasmine Begum. Crop recommendation with iot and ml. In *2023 4th International Conference for Emerging Technology (INCET)*, pages 1–5, 2023.

[7] SHILPA MANGESH PANDE, PREM KUMAR RAMESH, ANMOL ANMOL, B. R AISHWARYA, KARUNA ROHILLA, and KUMAR SHAURYA. Crop recommender system using machine learning approach. In *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, pages 1066–1071, 2021.