A MINI PROJECT

REPORT ON

SMART CROP PREDICTION SYSTEM

BACHELOR OF TECHNOLOGY

In

Computer Science & Engineering (DS)

Submitted By

Ayush Kumar(2201921540046)
Deepak Jaiswal(2201921540053)
Ankur Sanjay Yadav(2201921540032)
Akash Kumar Tiwari(2201921540017)

Under the Supervision of Ms. Roobal Yadav (Assistant Professor)

G.L. BAJAJ INSTITUTE OF TECHNOLOGY & MANAGEMENT, GREATER NOIDA

Affiliated to

DR. APJ ABDUL KALAM TECHNICAL UNIVERSITY, LUCKNOW

2024-2025

Declaration

We here with declare that the project work conferred during this report entitled "Smart Crop Prediction System" in partial fulfillment of the necessity for the award of the degree of Bachelor of Technology in Computer science and Data Science submitted to Dr A.P.J. Abdul Kalam Technical University, Uttar Pradesh, is an authentic record of our own work distributed in the Department of Computer science and Data Science G.L. Bajaj Institute of Technology & Management, Greater Noida. It contains no material antecedently printed or written by another person except wherever due acknowledgement has been created within the text. The project work reported during this report has not been submitted by us for award of the other degree or certification.

Signature:
Name: Ayush Kumar
Roll No: 2201921540046

Signature:
Name: Deepak Jaiswal
Roll No: 2201921540053

Signature:
Name: Ankur Sanjay Yadav
Roll No: 2201921540032

Signature:
Name: Akash Kumar Tiwari
Roll No: 2201921540017

Date:
Place: Greater Noida

Certificate

This is to certify that Project Report entitled "Smart Crop Prediction System" that is submitted by **Ayush Kumar, Deepak Jaiswal, Ankur Sanjay Yadav, Akash Kumar Tiwari** in partial fulfillment of the necessity for the award of degree B. Tech. in Department of Computer science and Data Science, G.L. Bajaj Institute of Technology & Management, affiliated by, Dr. A.P.J. Abdul Kalam Technical University, are records of the candidate's own work distributed by him below my/our oversight. The matter embodied during this thesis is original and has not been submitted for the award of the other degree.

Date:

Ms. Roobal Yadav
(Assistant Professor & Supervisor)

Dr. Mayank Singh
(Professor & Head of Department)

Acknowledgement

We would like to express our sincere thanks to our project supervisor Ms. Roobal Yadav and our Head of Department, Dr. Mayank Singh, for their invaluable guidance and suggestions. This project helped us to understand the concepts of machine learning, data analysis, and the application of algorithms like the Random Forest Classifier in agriculture. It also provided us hands-on experience with technologies such as Python, pandas, NumPy, scikit-learn, and data visualization tools. This project enriched our knowledge and experience of working in a team on a real-world problem. Additionally, we would like to express our gratitude to Ms. Roobal Yadav for her assistance in the preparation and review of our project.

Lastly, we would like to thank all the faculty members for their valuable time and support, which was instrumental in helping us successfully complete this project.

Names of Students

Ayush Kumar (2201921540046)

Deepak Jaiswal (2201921540053)

Ankur Sanjay Yadav (2201921540032)

Akash Kumar Tiwari (2201921540017)

ABSTRACT

Agriculture, a key component of economies worldwide, is currently facing numerous complex challenges such as resource scarcity, climate change, and the growing need to feed an increasing global population. Traditional farming practices, which often rely on farmers' instincts and experience, are becoming inadequate to meet these evolving demands. To address these issues, this project aims to develop a *Smart Crop Prediction System* using machine learning techniques, specifically the Random Forest Classifier. The system analyzes a variety of environmental and soil factors, including temperature, humidity, rainfall, and essential soil nutrients such as potassium, phosphorus, and nitrogen, to provide farmers with insightful and data-driven recommendations on optimal crop selection.

The system was built using Python and leverages key libraries such as pandas, NumPy, scikit-learn, and matplotlib for data preprocessing, analysis, and model training. Through the integration of machine learning and comprehensive data analysis, the system delivers highly accurate crop predictions, with an accuracy rate of 99.3%, confirmed through feature importance analysis, classification reports, and confusion matrices. This high level of accuracy ensures that farmers receive reliable, actionable insights that can guide their decisions, leading to optimized yields and promoting sustainable farming practices.

Looking ahead, future research will aim to enhance the system by developing user-friendly interfaces to make it more accessible to farmers, enabling them to effortlessly retrieve crop predictions. Moreover, real-world validation will be carried out using dynamic, real-time data, ensuring the system's adaptability and scalability in diverse agricultural settings. This Smart Crop Prediction System has the potential to revolutionize modern agriculture by blending advanced technological tools with traditional farming methods, providing scalable, data-driven solutions to key challenges such as resource efficiency, climate variability, and food security. Ultimately, it paves the way toward more resilient and sustainable agricultural practices.

TABLE OF CONTENT

LIST OF FIGURES

Chapter 1

Introduction

Overview

The best crops for an area are predicted by a Smart Crop Prediction System using machine learning and variables including soil type, weather, and market trends. By offering data-driven insights, lowering risks, and improving agricultural practices' efficiency, this technology assists farmers in maximizing sustainability and production.

1.1 Background

Despite its importance to the world economy, agriculture confronts several obstacles, including shifting market needs, soil erosion, and unpredictable weather patterns. Farmers have always chosen crops based on past knowledge and experience. But traditional approaches frequently fall short in tackling today's problems. Precision farming is becoming more and more possible because to machine learning, which uses massive datasets to enhance decision-making and produce precise forecasts.

Chapter 2

Literature Review

2.1 Evolution of Crop Prediction

From simple statistical models to sophisticated machine learning systems, crop prediction has changed throughout time. Early approaches concentrated on a small number of variables, such temperature and rainfall. Modern methods now incorporate a variety of variables, such as market trends, soil health measures, and data from remote sensing, thanks to the development of computing technology.

2.2 Effectiveness and Limitations

Machine learning algorithms have demonstrated efficacy in managing huge datasets and improving prediction accuracy. Limitations including the lack of data in rural areas, the high processing demands, and the interpretability of the model, however, continue to be major obstacles.

2.3 Technological Innovations in Agriculture

New technologies like drones, satellite images, and Internet of Things (IoT) devices have improved the amount and quality of agricultural data. When paired with machine learning, these technologies provide a strong foundation.

2.4 Ethical and Social Implications

Although there are many advantages to adopting technology, there are also drawbacks, including uneven access to resources, privacy problems with data, and the possibility of becoming overly dependent on automated processes. For agricultural growth to be inclusive and sustainable, ethical execution is essential.

Chapter 3

Existing Technologies

3.1. Python

- **Purpose:** Python serves as a high-level programming language for the machine learning algorithms, data processing, and model training.
- **Features:**

  Facilitates effective data preparation and model creation by supporting libraries such as scikit-learn, pandas, and NumPy.

  Incorporates machine learning methods like the Random Forest Classifier with ease.

  Allows for the display of data using Matplotlib and Seaborn to get insights into data patterns and feature relevance.

- **Impact:** Makes it easier to apply machine learning models and analyze data, which helps to provide accurate crop projections.

3.2. Random Forest Classifier

- **Purpose**: The purpose of this machine learning system is to forecast the best crop depending on environmental conditions.
- **Features:**

  For precise predictions, an ensemble approach that constructs several decision trees and aggregates their results.

  Addresses incomplete information and non-linear correlations between environmental factors.

  Explains the impact of variables including rainfall, temperature, and soil nutrients using feature significance analysis.

- **Impact**: Provides dependable data-driven insights to farmers, guaranteeing excellent crop prediction accuracy.

3.3. Scikit-learn

- **Purpose:** The system's classification methods are implemented using a machine learning library.
- **Features:**

Provides a range of machine learning models, such as Random Forest, facilitating rapid testing and training.

Offers features such as confusion matrices and classification reports for model assessment.

- **Impact:** Enhances the crop prediction model's performance and accuracy by facilitating its effective development and assessment.

3.4. Pandas and NumPy

- **Purpose:** Crucial libraries for the Smart Crop Prediction System's data processing and analysis.
- **Features:**

Pandas provides robust tools for organizing, cleaning, and managing big datasets in agriculture.
NumPy offers quick matrix operations and array calculations that are essential for handling environmental elements.
Impact: Makes it easier for the system to handle and prepare data for more precise model predictions.

3.5. Matplotlib and Seaborn

- **Purpose:** The goal of these visualization packages is to show data patterns and the connections between crop kinds and environmental conditions.
- **Features:**

Matplotlib generates simple plots to depict data distributions, including bar charts and histograms.
Pair plot matrices are among the sophisticated visualizations that Seaborn provides for gaining a deeper understanding of variable correlations.
Impact: Offers consumers lucid and educational visual outputs that aid in their comprehension of the facts and the system's decision-making procedure.

3.6. Pandas and NumPy

- **Purpose:** Crucial libraries for the Smart Crop Prediction System's data processing and analysis.

- **Features:**
  - Pandas provides robust tools for organizing, cleaning, and managing big datasets in agriculture.
  - **Impact:** Makes it easier for the system to handle and prepare data for more precise model predictions.

3.7. Matplotlib and Seaborn

- **Purpose:** The goal of these visualization packages is to show data patterns and the connections between crop kinds and environmental conditions.
- **Features:**

Matplotlib generates simple plots to depict data distributions, including bar charts and histograms.

Pair plot matrices are among the sophisticated visualizations that Seaborn provides for gaining a deeper understanding of variable correlations.

- **Impact:** Offers consumers lucid and educational visual outputs that aid in their comprehension of the facts and the system's decision-making procedure.

3.8. Jupyter Notebooks

- **Purpose:** The goal of this open-source online application is to create, record, and distribute data analysis processes and machine learning models.
- **Features:**

Integrates visuals, rich text, and live code execution into a one page.

Permits code to be tested and adjusted interactively while the crop prediction system is being developed.

Makes it simple to create reports and record results.

- **Impact:** Facilitates effective machine learning process presentation and prototyping by improving transparency and cooperation.

3.9. Seaborn

- **Purpose:** A Matplotlib-based Python data visualization framework for producing visually appealing and educational statistics visuals.
- **Features:**
  - High-level APIs for creating informative graphs, including pair plots, scatter plots, and heatmaps, are provided.
  - Illustrates how various environmental factors (such as temperature, precipitation, and soil nutrients) relate to one another and how these factors affect crop forecast.
- **Impact**: By producing understandable and useful representations of intricate data linkages, the system becomes easier to interpret.

3.10. Scikit - learn’s Grid Search CV

- **Purpose:** This module's function is to adjust the Random Forest Classifier's and other machine learning models' hyperparameters.
- **Features**:

Tests various combinations of hyperparameters to automatically find the optimal model parameters.

Improves model performance by optimizing important parameters such as the maximum depth of each tree or the number of trees in the forest.

- **Impact:** Chooses the best Random Forest model parameters to increase the system's prediction accuracy.

3.11. GitHub

- **Purpose:** The goal of the web-based platform is to facilitate collaborative development and version control,
- **Features:**
  - The team can manage code changes and roll back to previous versions if needed thanks to the feature that allows version tracking.
  - Enables teamwork via code reviews, pull requests, and shared repositories.

Integrates for more efficient development with cloud-based technologies such as Google Colab.

- **Impact:** Promotes better teamwork and project management, guaranteeing systematic and uniform code development.

Chapter 4

Purpose and Significance

4.1 Purpose: Giving farmers and other agricultural stakeholders useful insights from data analytics and machine learning is the main goal of the Smart Crop Prediction System. In particular, the system seeks to:

- Improve Crop Selection: Give farmers advice on which crops to grow based on market demands, soil properties, and weather patterns.
- Enhance Productivity: Reduce crop failure risks to boost agricultural productivity and efficiency.
- Optimize Resources: By focusing on the demands of certain crops, recommendations may be made to encourage the effective use of resources like water, fertilizer, and pesticides.
- Support Sustainability: Promote ecologically responsible agricultural methods by including conservation and climate adaption techniques into crop choices.
- Reduce Economic Risks: Assist farmers make well-informed decisions by offering accurate forecasts and market information to assist them minimize financial losses.
- Promote Technology Adoption: Close the gap between conventional agricultural practices and contemporary technical developments to enable everyone to use precision agriculture.

4.2 Significance: The capacity of the Smart Crop Prediction System to tackle important issues in contemporary agriculture while promoting sustainability and global food security makes it significant. Its influence is highlighted by the points below:

- **Economic Benefits:** By empowering farmers to select crops that maximize earnings and minimizing resource waste, the system immediately raises farmer incomes and lowers agricultural operating expenses.
- **Environmental Impact:** By reducing excessive use of pesticides, fertilizers, and water, precision agriculture helps to conserve natural resources and reduce pollution in the environment.
- **Adaptation to Climate Change:** By identifying trends and forecasting which crops are most appropriate for certain areas under various climatic circumstances, the system assists farmers in adjusting to the changing environment.
- **Enhanced Food Security:** The method helps to alleviate the world's food demand issues and lessen hunger by increasing production and guaranteeing a consistent supply of crops.
- **Technological Advancement in Agriculture:** This project demonstrates the potential of innovation in conventional industries by incorporating state-of-the-art technologies like data analytics, IoT, and machine learning into farming.
- **Reduction in Crop Failure:** By enabling preventive interventions through real-time data integration and predictive analysis, the chance of crop failure brought on by unanticipated variables like insect infestations or extreme weather occurrences is decreased.
- **Empowerment of Smallholder Farmers:** The system democratizes access to technology by offering small and marginal farmers, who are frequently left out of the advantages of expensive agricultural research, scalable and reasonably priced solutions.
- **Policy and Planning Support:** Using the information produced by the system, governments and agricultural associations may plan agricultural operations at the regional or national level, distribute resources, and create efficient policies.
- **Global Adaptability:** The system may be used in different parts of the world and can help farmers all over the world because of its modular architecture and dependence .

Chapter 5

Key Features

5.1 Data Collection and Preprocessing

Market pricing, historical temperature data, and assessments on soil health are just a few of the many datasets that the system incorporates. By encoding categorical data, managing missing values, and normalizing parameters, preprocessing guarantees the quality of the data.

5.2 Machine Learning Algorithms

The system uses neural networks, Random Forest, and Support Vector Machines (SVM) to predict crops. The scalability, accuracy, and precision of each model are assessed.

5.3 Scalability

Applications from tiny farms to vast agricultural regions can benefit from the architecture's ability to handle massive amounts of data.

5.4 User-Friendly Interface

With a user-centric design, farmers can enter basic information like location and soil type and get recommendations that are simple to grasp.

5.5 Real-Time Updates

Weather APIs and IoT devices are integrated to automatically update weather forecasts to reflect current conditions.

5.6 Evaluation Metrics

Metrics like R-squared, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) are used to evaluate the performance of the model.

5.7 Dataset Utilization

The FAO's agriculture database and NASA's climate data are two examples of publicly available datasets that are used to train and test the system.

5.8 Accessibility Features

In addition to offline capabilities for areas with poor internet access, the system supports many languages.

5.9 Innovative Enhancements

Crop image-based disease prediction and suggestions for sustainable agricultural methods are further features.

Chapter 6

Overview of the System Design

6.1 SYSTEM DESIGN

The system is made up of a number of parts, including user interaction, predictive modeling, and data collection. Through a centralized processing unit, these parts are connected.

6.1.1 Information Gathering

Sensors, public databases, and user input are the sources of data.

6.1.2 Training of Models

Supervised learning approaches are used to train machine learning models on historical data. Performance is optimized through feature selection and hyperparameter adjustment.

6.1.3 Module for Prediction

Based on processed data, the model makes suggestions while taking environmental factors, market trends, and crop appropriateness into account.

6.1.4 Illustration

Graphs and heatmaps are used in interactive dashboards to show trends, insights, and forecast outcomes.

6.1.5 Features of Security

Strong encryption guarantees the privacy and security of user data.

6.2 Problem Description

Crop selection inefficiencies and losses result from farmers' lack of access to precise and useful technologies.

6.3 Goal

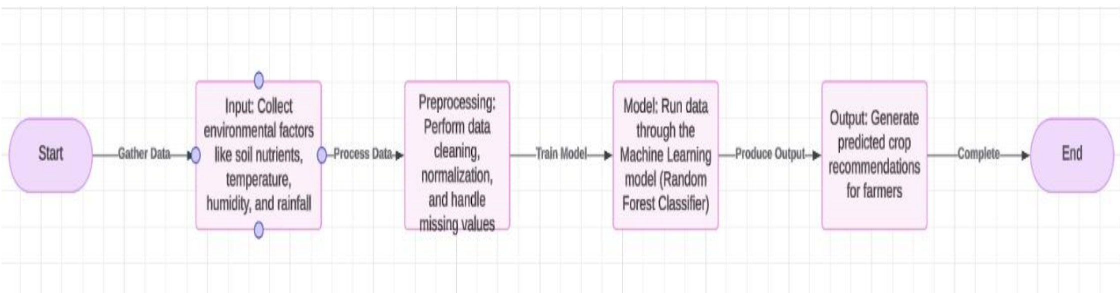
The goal is to offer a scalable and reasonably priced system for data-driven crop selection.

6.4 Major Issue Solved

The method tackles problems including resource waste, crop failure brought on by inadequate planning, and restricted access to market intelligence.

6.5 Anticipated Result

It is anticipated that the system will boost farmer earnings, decrease its negative effects on the environment, and increase crop efficiency.



Chapter 7

System Architecture

7.1 Layer of Input

Users use a mobile or web application to enter their location, soil parameters, and other preferences.

7.2 Layer of Data Processing

In order to ensure interoperability with machine learning techniques, this layer preprocesses and normalizes input.

7.3 Forecast Model

With the use of market and meteorological data, pre-trained computers evaluate inputs and suggest crops.

7.4 Layer of Post-Processing

Improves the interpretability and usefulness of model results.

7.5 Output Layer

Presents useful information in the form of interactive maps, graphs, and text.

7.6 Module for Continuous Learning

Gradually increases prediction accuracy by adjusting to fresh data and human input.

7.7 Layer of Development

Has a modular architecture that makes it simple to include additional datasets and functionality.

7.8 Supplementary Elements

Includes APIs for remote sensing data, soil testing kits, and weather updates.

Chapter 8

Future Scope

8.1 Improved Predictive Precision

Use sophisticated deep learning methods for increased accuracy, such as transformers and LSTMs.

8.2 Integration of Multimodal Data

Using satellite photos and drone visual data to supplement numerical data.

8.3 Adaptability Worldwide

Increasing the system's support for various agricultural regions across the world in terms of both dataset and algorithm.

8.4 Ecological Farming Methods

Incorporating sustainability criteria into forecasts to help farmers adopt environmentally responsible practices.

8.5 Integration with Emerging Technologies

Using big data analytics for trend forecasting and blockchain for supply chain traceability.

8.6 Support for Multiple Languages

Creating interfaces in several regional languages will guarantee that farmers in various areas can access them.

8.7 Social and Ethical Aspects

Ensuring that the use of technology doesn't worsen environmental damage or inequality.

Chapter 9

Result & Conclusions

The Smart Crop Prediction System serves as an example of how data analytics and machine learning can revolutionize contemporary agriculture. Through the integration of technical innovations with pragmatic farming requirements, the system tackles important issues like crop failure, market instability, and resource inefficiency. Farmers, legislators, and society at large stand to gain much from its implementation.

Key Insights:

- Improved Ability to Make Decisions: Farmers are empowered by the system's precise, data-driven recommendations, which help them choose crops and allocate resources wisely.

- Sustainability: The system supports effective use of herbicides, fertilizers, and water, which is in line with international initiatives for environmentally friendly farming practices.

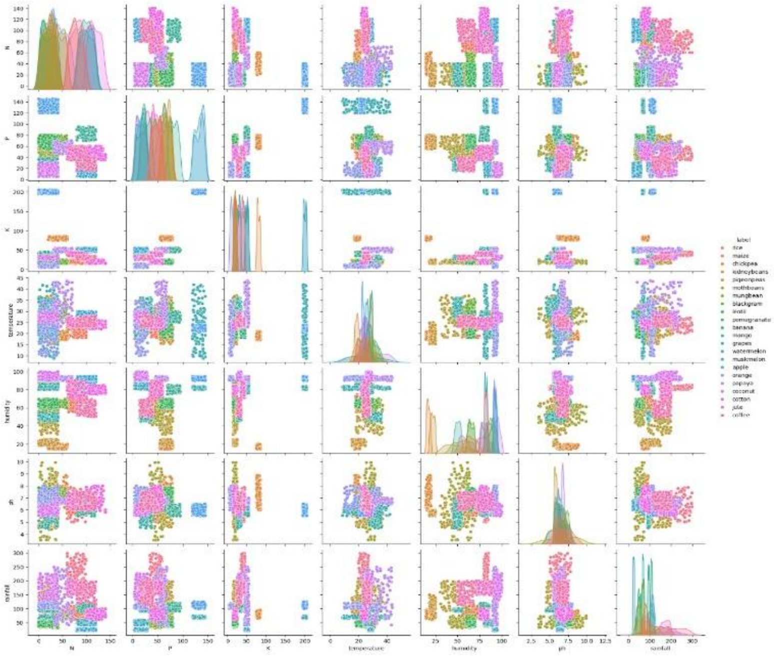
- Economic Growth: Farmers benefit financially from higher productivity and lower losses, which promotes economic stability in farming communities.

- Adaptability: The system is a worldwide feasible solution due to its modular and scalable architecture, which guarantees its application across a variety of geographic locations and farming practices.

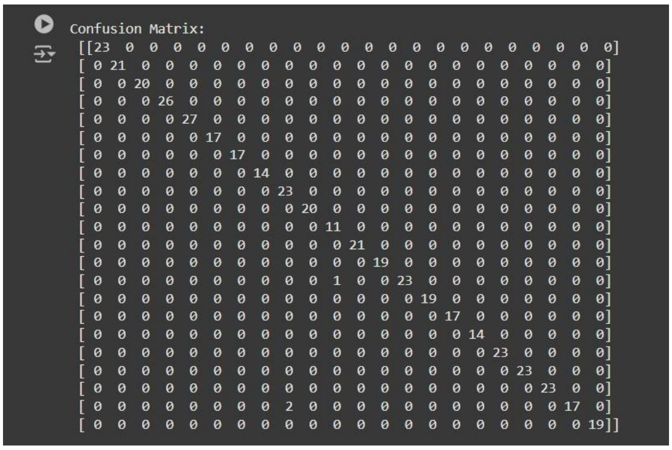
- Innovation in Agriculture: The initiative establishes a standard for upcoming developments in the field by bridging the gap between conventional farming methods and state-of-the-art technology.

  The following are pair plot matrix, confusion matrix and bar chart of the model.

1. Pair Plot Matrix :- It is used to visualize the relationships between different environmental factors (such as Nitrogen, Phosphorus, and Rainfall) and how they relate to different crop types.

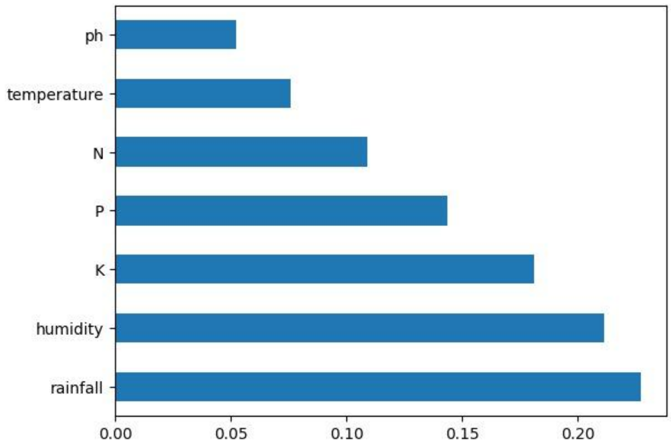


2. Confusion Matrix :- It is used to show performance of Random Forest Classifier.



```
Confusion Matrix:
[[23  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0 21  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0 20  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0 26  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0 27  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0 17  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0 17  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0 14  0  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0 23  0  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0 20  0  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0 11  0  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0 21  0  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0 19  0  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  1  0  0 23  0  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0 19  0  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 17  0  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 14  0  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 23  0  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 23  0  0  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 23  0  0]
 [ 0  0  0  0  0  0  0  0  2  0  0  0  0  0  0  0  0  0  0  0 17  0]
 [ 0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0 19]]
```

3. Bar Chart :- It visualizes the importance of various environmental features like nitrogen, phosphorus, potassium, humidity and rainfall.



Final Thoughts:

Result & Conclusion: The Smart Crop Prediction System is a step toward transforming agriculture as a resilient and sustainable sector, not only a technical tool. Precision farming has advanced significantly as a result of its capacity to combine market trends, soil health, and weather patterns into useful data.

Even while the system has a significant influence currently, there is a ton of room for innovation. The system may become even more efficient with future improvements including real-time analytics, sophisticated AI methods, and wider dataset integration. Its broad adoption will also depend on resolving issues like smallholder farmers' accessibility and guaranteeing ethical data use.

To sum up, the Smart Crop Prediction System is an essential instrument for tackling the worldwide issues of economic inequality, climate change, and food security. The system has the potential to significantly contribute to the development of a sustainable and profitable agricultural future if it keeps innovating and developing

References

[1] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324

[2] Chandio, I. A., Zhang, F., & Wang, G. (2018). Crop yield prediction using machine learning: A case study of winter wheat yield. *Computers and Electronics in Agriculture*, 151, 61-72. https://doi.org/10.1016/j.compag.2018.05.012

[3] Fang, Y., Fan, J., & Li, Z. (2019). Data-driven crop yield prediction using ensemble learning techniques. *International Journal of Agricultural and Biological Engineering*, 12(2), 40-48. https://doi.org/10.25165/j.ijabe.20191202.5060

[4] Khalil, U., & Naeem, M. (2021). An intelligent crop recommendation system based on soil and environmental parameters using machine learning. *Sustainability*, 13(10), 5673. https://doi.org/10.3390/su13105673

[5] Mohapatra, A., Nayak, D. R., Tripathy, A. K., & Pattnaik, P. K. (2020). A machine learning approach for prediction of crop yield based on climate and soil parameters. *Journal of King Saud University - Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2020.01.010

[6] Shahhosseini, M., Hu, G., & Archontoulis, S. V. (2019). Forecasting corn yield with machine learning ensembles. *Agricultural & Forest Meteorology*, 266, 367-380. https://doi.org/10.1016/j.agrformet.2018.09.017

[7] Huang, Y., Lan, Y., Thomson, S. J., Fang, A., & Hoffmann, W. C. (2010). Development of soft computing and applications in

agricultural and biological engineering. *Computers and Electronics in Agriculture*, 71(1), 107-127. https://doi.org/10.1016/j.compag.2010.09.001

[8] Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data mining: Practical machine learning tools and techniques* (4th ed.). Morgan Kaufmann.

[9] Wu, D., Li, M., Liu, Z., & Yu, J. (2020). Crop yield prediction based on machine learning: A comparative study. *Sustainability*, 12(9), 3427. https://doi.org/10.3390/su12093427

[10] Zhang, X., Zhang, L., & Wei, C. (2018). Feature importance analysis in agricultural data using random forest. *Journal of Agriculture and Technology*, 13(4), 200-208.

[11] Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. *Frontiers in Plant Science*, 10, 621. https://doi.org/10.3389/fpls.2019.00621

[12] Adhikari, K., & Singh, V. P. (2019). Evaluation of machine learning methods for crop yield estimation: A case study on maize yield prediction. *Field Crops Research*, 234, 12-22. https://doi.org/10.1016/j.fcr.2019.01.014

[13] Ahmad, M., & Chandio, A. A. (2020). Application of machine learning algorithms for crop yield prediction. *Sustainability*, 12(6), 2204. https://doi.org/10.3390/su12062204

[14] Jeong, J., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., & Kim, S.-H. (2016). Random forests for global and regional crop yield predictions. *PLOS ONE*, 11(6), e0156571. https://doi.org/10.1371/journal.pone.0156571

[15] Lobell, D. B., Thau, D., Seifert, C., Engle, E., & Little, B. (2015). A scalable satellite-based crop yield mapper. *Remote Sensing of Environment*, 164, 324-333. https://doi.org/10.1016/j.rse.2015.04.021