DESIGN AND DEVELOPMENT OF CROP RECOMMENDATION SYSTEM USING MACHINE LEARNING

**A**

**MINOR PROJECT-I REPORT**

Submitted in partial fulfillment of the requirements for the degree of

## BACHELOR OF TECHNOLOGY

in

## CSE-ARTIFICIAL INTELLIGENCE & DATA SCIENCE

By

## GROUP NO. 03

|  |  |
| --- | --- |
| Aman Sahu | 0187AS221007 |
| Shyam Sahu | 0187AS221059 |
|  |  |

Under the guidance of

## Prof. Abhuday Tripathi

(ASSISTANT PROFESSOR)

## Department of CSE-Artificial Intelligence & Data Science Sagar Institute of Science & Technology (SISTec) Bhopal (M.P.)

### Approved by AICTE, New Delhi & Govt. of M.P.

**Affiliated to Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal (M.P.)**

## DECEMBER-2024

**Sagar Institute of Science & Technology (SISTec), Bhopal**

Department of CSE-Artificial Intelligence & Data Science



CERTIFICATE

I hereby certify that the work which is being presented in the B.Tech. Minor Project-I Report entitled **Design and Development of Crop Recommendation System,** in partial fulfillment of the requirements for the award of the degree of ***Bachelor of Technology***, submitted to the Department of *CSE* with *Artificial Intelligence & Data Science*, *Sagar Institute of Science & Technology (SISTec)****,*** Bhopal (M.P.) is an authentic record of my own work carried out during the period from July-2024 to Dec-2024 under the supervision of **Prof. Abhuday Tripathi (Assistant Professor)**.

The content presented in this project has not been submitted by me for the award of any other degree elsewhere.

*Signature*

|  |  |  |
| --- | --- | --- |
| ***Aman Sahu*** |  | ***Shyam Sahu*** |
| ***0187AS221007*** |  | ***0187AS221059*** |

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

***Date:***

### Prof. Abhuday Tripathi (Assistant Professor)

**Dr. Vasima Khan HOD, CSE-AI&DS**

**Dr. D.K. Rajoriya Principal, SISTec**

**ACKNOWLEGMENT**

We would like to express our sincere thanks to **Dr. D.K. Rajoriya, Principal, SISTec** and **Dr. Swati Saxena, Vice Principal, SISTec** Gandhi Nagar, Bhopal for giving us an opportunity to undertake this project.

We also take this opportunity to express a deep sense of gratitude to **Dr. Vasima Khan, HOD**, **Department of CSE-Artificial Intelligence & Data Science** for her kindhearted support.

We extend our sincere and heartfelt thanks to our Guide, **Prof. Abhuday Tripathi** for providing us with the right guidance and advice at the crucial junctures and for showing us the right way.

I am thankful to the **Project Coordinator, Ms. Ruchi Jain** who devoted her precious time in giving us the information about the various aspect and gave support and guidance at every point of time. I am thankful to their kind and supportive nature. His inspiring nature has always made my work easy.

I would like to thank all those people who helped me directly or indirectly to complete my project whenever I found myself in any issue.

**TABLE OF CONTENTS**

|  |  |  |  |
| --- | --- | --- | --- |
| **TITLE** |  |  | **PAGE NO.** |
| Abstract |  |  | i |
| List of Abbreviation | |  | ii |
| List of Figures | |  | iii |
| Chapter 1 | Introduction | | 1 |
|  | 1.1 | About Project | 2 |
|  | 1.2 | Project Objectives | 2 |
| Chapter 2 | Software & Hardware Requirements | | 5 |
| Chapter 3 | Problem Description | | 8 |
| Chapter 4 | Literature Survey | | 11 |
| Chapter 5 | Software Requirements Specification | | 13 |
|  | 5.1 | Functional Requirements | 13 |
|  | 5.2 | Non-Functional Requirements | 15 |
| Chapter 6 | Software Design | | 18 |
|  | 6.1 | Use Case Diagram | 19 |
|  | 6.2 | Data Flow Chart | 20 |
| Chapter 7 | Machine Learning Module | | 21 |
|  | 7.1 Data set Description | | 22 |
|  | 7.2 Pre-processing Steps | | 22 |
|  | 7.3 Data Visualization | | 23 |
|  | 7.4 ML Model Description | | 25 |
|  | 7.5 Model Workflow | | 25 |
| Chapter 8 | Front End Connectivity | | 26 |
| Chapter 9 | Future work | | 28 |
| Chapter 10 | Coding |  | 31 |
| Chapter 11 | Output Screens | | 34 |
| References |  |  | 37 |
| Appendix-1: Glossary of Terms | | | 39 |
| Appendix-2 Project Summary | | | 41 |

SISTec/BTech/AD/2024/5/Minor Project\_I/03

**ABSTRACT**

Crop Recommendation System for agriculture is based on various input parameters. This proposes a hybrid model for recommending crops to Indian states by considering various attributes such as potassium, Nitrogen, phosphorus, Rainfall, Temperature, Fertilizers, Pesticides and season.

The recommender model is built as a hybrid model using the classifier machine learning algorithm. Based on the appropriate parameters, the system will recommend the crop.

Technology based crop recommendation system for agriculture helps the farmers to increase the crop yield by recommending a suitable crop for their land with the help of geographic and the climatic parameters.

The proposed hybrid recommender model is found to be effective in recommending a suitable crop. Crop yield production value updation has a positive practical significance for guiding agricultural production and for notifying the change in market rate of crop to the farmer.

The concept of this paper is to implement the crop selection method so that this method helps in solving many agriculture and farmers problems. This improves our Indian economy by maximizing the yield rate of crop production. Different types of land condition. So the quality of the crops are identified using ranking process. By this process the rate of the low quality and high quality crop is also notified.

The usage of ensemble of classifiers paves a path way to make a better decision on predictions due to the usage of multiple classifiers. Further, a ranking process is applied for decision making in order to select the classifiers results.

SISTec/BTech/AD/2024/5/Minor Project\_I/03

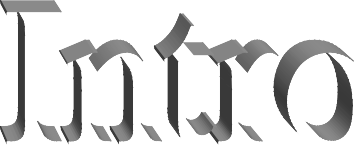
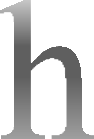
**LIST OF ABBREVIATIONS**

|  |  |
| --- | --- |
| **ACRONYM** | **FULL FORM** |
| ML | Machine Learming |
| HTML | Hyper Text Markup Language |
| CSS | Cascading Style Sheets |
| JS | Java Script |

SISTec/BTech/AD/2024/5/Minor Project\_I/03

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **FIG. NO.** | **TITLE** | **PAGE NO.** |
| 6.1 | Use case | 20 |
| 6.2 | Data Flow | 21 |
| 7.1 | Pair Plot | 24 |
| 7.2 | Confusion Matrix | 25 |
| 7.3 | Prediction curve | 25 |
| 7.4 | Model Workflow | 26 |



**CHAPTER-1**

**INTRODUCTION**

## ABOUT PROJECT

A Crop Recommendation System is an intelligent system designed to help farmers , agriculture experts, or even governments make informed decisions about which crops to cultivate in a particular location. This decision-making process is based on various parameters such as soil quality, weather conditions, geographical location, and economic factors. By leveraging data and advanced technologies, these systems aim to optimize agricultural productivity and sustainability.

## PROJECT OBJECTIVES

**Optimize Crop Selection**: Recommending the most suitable crops for a specific location based on factors such as soil type, climatic conditions, water availability, and market trends. This ensures that farmers choose crops that have the highest potential for growth, reducing the chances of crop failure.

**Increase Agricultural Productivity**: By providing scientific and data-driven recommendations, the system helps maximize crop yields. Efficient use of resources like fertilizers, water, and land improves overall agricultural output.

**Data-Driven Decision Making**: Integrating historical data, real-time weather updates, and soil analyses helps farmers make informed decisions.

**Efficient Resource Utilization**: By suggesting crops that match the soil's nutrient profile, the system minimizes the need for external inputs. Reduces wastage of water, fertilizers, and pesticides, lowering farming costs.

**Validation and Feedback:** Positive feedback from users who noted measurable improvements in productivity and income.Recommendations to enhance the system by including additional parameters like pest outbreaks, real-time weather forecasts, and market price predictions.

**Documentation and Knowledge Sharing**: Document the project findings, methodologies, and outcomes to facilitate knowledge sharing and future research at the intersection of crop recommendation and Data Science.

**Evaluation of Impact**: Evaluate the impact of adopting recommendation and Data Science techniques on detection accuracy, system quality, and user satisfaction.

Through these objectives, the project aims to demonstrate the feasibility and benefits of combining Data Science practices with recommendation techniques for agriculture. It seeks to provide researchers and practitioners with methodologies to develop more robust, adaptable, and accurate crop recommendation systems.

The complexity in agricultural decision-making arises from diverse soil types, fluctuating weather conditions, and the intricacies of real-world farming scenarios. Traditional approaches to crop selection have often struggled to account for such complexities, leading to solutions that fall short of precision or adaptability.

Data-driven recommendations offer a paradigm shift by advocating for an in-depth understanding of environmental and market factors, facilitating more precise identification of suitable crops. By focusing on the unique characteristics of soil, climate, and regional demands, such systems enable the development of models that more accurately reflect the realities of farming.

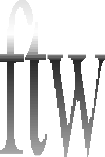
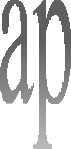
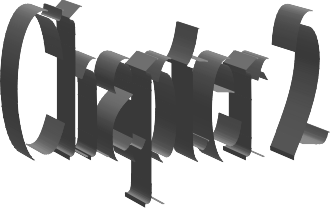
Simultaneously, the rise of Data Science has transformed the way agricultural data is analyzed to optimize decision-making. Techniques such as machine learning, statistical analysis, and predictive modeling have become crucial for achieving high accuracy in crop recommendations and for tailoring solutions to diverse farming conditions.

Recognizing the complementary nature of recommendation systems and Data Science, this project seeks to leverage the strengths of both domains to address the challenges of modern agriculture. By integrating data-driven insights with evidence-based methodologies, the project aims to create recommendation models that are not only technically robust but also aligned with the practical needs and complexities of farmers.

Moreover, this project is driven by a commitment to empirical validation and continuous improvement. Through rigorous testing, farmer feedback, and iterative refinement, we seek to ensure that the resulting recommendation models meet the highest standards of quality, usability, and accuracy.

In documenting our findings, methodologies, and outcomes, we aim to contribute to the expanding body of knowledge at the intersection of agriculture and Data Science. By sharing our experiences and insights, we hope to inspire further research and innovation in this promising area.

Ultimately, the success of this project will be measured not only by the tangible outcomes achieved, such as increased productivity and farmer income, but also by the impact it has on the broader field of agricultural technology. By demonstrating the feasibility and benefits of incorporating Data Science into crop recommendation systems, we aim to equip researchers and practitioners with tools and methodologies to tackle the challenges of modern farming with precision and confidence.



**CHAPTER-2**

**SOFTWARE AND HARDWARE SPECIFICTAION**

This chapter outlines the essential software and hardware components required for the successful implementation and deployment of the Crop Recommendation System.

## SOFTWARE REQUIREMENTS:

**Integrated Development Environment (IDE**): Utilize tools like Visual studio or Jupyter Notebook for efficient code writing, testing, and debugging of crop recommendation algorithms.

**Version Control System (VCS):** Employ Git to manage project versions and facilitate team collaboration, ensuring traceability of changes.

**Programming Languages:** Python for its flexibility in data preprocessing and machine learning, and Sqlite3 for managing any related data storage.

**Data Analysis Libraries**: Pandas and NumPy for data manipulation, and Scikit-learn for building model related to crop recommendation system.

**Machine Learning and Deep Learning Frameworks**: Scikit-learn for building and training recommendation models, with a focus on achieving high accuracy in crop recommendation system.

**Data Visualization Tools**: Matplotlib and Seaborn for visualizing data distributions and analysis results.

**Database** Management Systems (DBMS): Sqlite3 for reliable storage and retrieval of segmented images or related data.

**Text Editors**: Visual Studio Code or Sublime Text for quick code adjustments and script editing

**Virtual Environment Management:** Anaconda or Virtualenv for creating isolated environments to manage project dependencies effectively.

## HARDWARE REQUIREMENTS:

**High-Performance GPU**: A dedicated GPU, such as NVIDIA GTX 1080 or better, for accelerated training and processing of machine learning models for recommendation.

**Workstation/Desktop/Laptop**: A powerful workstation or laptop with a high-performance CPU, at least 4GB of RAM, and SSD storage for efficient development and training tasks.

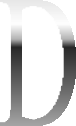
**Networking Equipment**: Router and Ethernet cables for stable network connectivity, particularly for remote collaboration or data transfer.

**Input/Output Devices**: Keyboard, mouse, and monitor for user interaction with computing resources during model training and testing phases.

**Power Backup**: Uninterruptible Power Supply (UPS) to prevent data loss and safeguard ongoing training sessions during power outages.

**Cooling Systems**: Adequate cooling, such as fans or liquid cooling systems, to maintain optimal temperatures, particularly during intensive GPU-based model training.

**Internet Connectivity**: Reliable broadband connection for recommend best suitable crop.



**CHAPTER-3**

**PROBLEM DESCRIPTION**

In this chapter, we delve into the specific issues within the domain of crop recommendation that our project aims to address. We provide a detailed description of the challenges faced in conventional methods of crop selection, such as reliance on traditional knowledge and generalized practices. Additionally, we highlight the need for an automated and efficient solution that leverages data-driven techniques like machine learning. This chapter sets the foundation for understanding how our system overcomes these challenges to optimize crop selection and improve agricultural outcomes.

## CURRENT SCENARIO AND CHALLENGES

In this section, we provide an overview of the existing landscape of crop recommendation methods used in agriculture. We highlight the limitations and shortcomings of traditional methods, such as manual decision-making by farmers or basic agricultural practices that often struggle with precision in selecting the most suitable crops for specific conditions. These approaches frequently lack accuracy, consistency, and scalability, making it difficult to achieve reliable results in diverse applications, including precision farming, resource optimization, and climate-resilient agriculture.

## IMPORTANCE OF AUTOMATION

Here, we emphasize the importance of automating the crop recommendation process to improve accuracy and efficiency. We discuss the potential impact of automated crop recommendations on various fields, underscoring the benefits of precise and consistent decision-making techniques. By automating the crop recommendation process, we aim to streamline agricultural workflows, enhance the quality of farming practices, and enable more advanced applications, such as precision farming and sustainable agricultural development.

## OBJECTIVES OF THE PROJECT

This section outlines the specific objectives that our project seeks to achieve in addressing the problem of crop recommendation. Our objectives include:

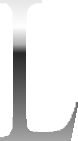
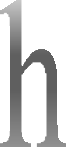
* Developing a robust and accurate crop recommendation system using advanced data analysis and machine learning algorithms.
* Leveraging machine learning models to analyze soil, weather, and environmental conditions to recommend the most suitable crops.
* Providing an adaptable framework for crop recommendations that can be applied across various agricultural regions and farming scenarios.
* Evaluating the effectiveness of the system through rigorous testing and validation on multiple datasets with varied soil types, climates, and crop patterns.

## SCOPE OF THE PROJECT

Here, we define the scope of our project by outlining the boundaries and limitations within which we will operate. We specify the types of images targeted by our hair detection system, such as close-up hair images or full portraits, and the environments in which the system is intended to function, including forensic labs and cosmetology applications. Additionally, we highlight any constraints, such as hardware requirements or dataset limitations, that may impact the implementation and deployment of the system.

## SUMMARY

To conclude the chapter, we summarize the key points discussed regarding the problem of crop selection, its significance in agriculture and related fields, and the objectives and scope of our project. This chapter highlights the critical challenges faced by farmers in choosing suitable crops and underscores the importance of a data-driven solution to address these issues. It sets the foundation for subsequent chapters, providing context and justification for the development of our automated crop recommendation system.



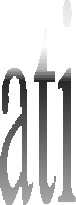
**CHAPTER-4**

**LITERATURE SURVEY**

The paper [1] states the requirements and planning needed for developing a software model for precision farming is discussed. It deeply analyses the basics of precision farming. The author’s start from the basics of precision farming and move towards developing a model that would support it. This paper describes a model that applies Precision Agriculture (PA) principles to small, open farms at the individual farmer and crop level, to affect a degree of control over variability. The comprehensive objective of the model is to deliver direct advisory services to even the smallest farmer at the level of his/her smallest plot of crop, using the most accessible technologies such as SMS and email. This model has been designed for the scenario in where the average holding size is much lower than most of India. Hence this model can be deployed elsewhere in India only with minor modifications. The paper [2] makes a comparative study of classification algorithms and their performance in yield prediction in precision agriculture. These algorithms are implemented in a data set collected for several years in crop recommend.. The paper [3] states the necessity for crop recommendation and its help in a nation’s strategic policy making in agriculture. A framework eXtensible Crop recommendation Framework (XCYPF) is developed. It facilitates flexible inclusion of various techniques towards crop recommendation. A tool was also developed that would help people to recommend crop for various dependant and independent variables.

The paper [4] states the usage of agricultural data with data mining and visual data mining techniques are depicted. This paper reduces the high dimensional agricultural data to smaller size to acquire useful knowledge related to crop, input application(like fertilizers).The techniques used is Self-organizing maps and multi-dimensional scaling techniques (Sammon’s mapping) to reduce the data. The conclusion derived is that Self-organizing maps is suitable when dataset is large and Sammon’s maping is suitable when data set is small. The paper [5] depicts the importance of crop selection and the factors deciding the crop selection like production rate, market price and government policies are discussed. This paper proposes a Crop Selection Method (CSM) which solves the crop selection problem. It suggests a series of crop to be selected over a season considering factors like weather, soil type, water density, crop type. The predicted value of influential parameters determines the accuracy of CSM. Hence there exists a need to include a prediction method with improved accuracy and performance. Data mining techniques in paper [6] are used to estimate the crop recommend for cereal crops in major districts of Bangladesh. The methodology comprises of two parts namely Clustering (for creating district clusters) and Classification using Logistic Regression,. The accuracy of prediction lies in the range of 90-95. The data set included 5 environmental variables, 3 biotic variables and 2 area related variables to determine the crop recommend in different districts. The paper proposed a future work of geospatial analysis to improve accuracy.

The paper [7] aims to solve the crucial problem of selecting the classifiers for the ensemble learning. A method to select a best classifier set from a pool of classifiers has been proposed. The proposal aims to achieve higher accuracy and performance. A method called SAD was proposed based on accuracy and classification performance. Using Q statistics, the dependency between most relevant and accurate classifiers is identified. The classifiers which were not chosen were combined to form the ensemble. This measure is supposed to ensure higher performance and diversity of the ensemble. Various methods such as SA (Selection by Accuracy), SAD (Selection by accuracy and Diversity) and NS (No selection) algorithm were identified. Finally it is inferred that SAD works better than others. The paper [8] proposes various classification methods to classify the liver disease data set. The paper emphasizes the need for accuracy because it depends on the dataset and the learning algorithm. Classification algorithms such as J48, Logistic regression correction rate among them. The performance of the models where compared with accuracy and computational time. It was concluded that all the classifiers showed improved predictive performance. Multilayer perceptron show the highest accuracy among the proposed algorithms. The paper [9] tries to solve the problem of food insecurity in Egypt. It proposes a framework which would predict the production, and import for that particular year. It uses At the end of the process we would be able to visualize the amount of production import, need and availability. Therefore it would help to make decisions on whether food has to be further imported or not. The soil datasets in paper [10] are analyzed and a category is predicted. From the predicted soil category the crop recommend is identified as a Classification rule. Logistic Regression algorithms are used for crop recommendation. The future work stated is to create efficient models using various classification techniques such as support vector machine, principal component analysis.



**CHAPTER-5**

**SOFTWARE REQUIREMENT AND SPECIFICATION**

## INTRODUCTION:

The Software Requirements Specification (SRS) outlines the functional and non-functional requirements of the Crop Recommendation System. This document serves as a guide for software developers, designers, and stakeholders involved in the development process.

## FUNCTIONAL REQUIREMENTS:

### Data Collection:

The system shall collect data from various sources such as soil testing results, weather data, and user inputs.

It should support multiple input formats, including numerical data (e.g., pH levels, temperature).

### Preprocessing:

The system shall preprocess input data to ensure consistency and accuracy by applying techniques like normalization, filtering, and missing data handling.

### Crop Suitability Recommendation:

The system shall recommend the most suitable crops based on soil conditions, climate, and water availability..

Recommendations should consider crop rotation, pest resistance, and market demand to enhance productivity and profitability.

### Feature Extraction:

Once soil and environmental regions are detected, the system shall extract specific features, such as soil nutrient levels, pH, moisture content, and temperature.

Feature extraction techniques such as remote sensing, soil sample analysis, or advanced machine learning models shall be employed.

### Classification and Analysis:

The system shall classify segmented agricultural regions based on specific attributes like soil type, moisture level, and nutrient content.  
Analysis of extracted features will enable further applications, such as precision farming, yield prediction, or sustainable resource management.

### Output and Visualization:

The system shall provide visualized outputs highlighting suitable crop regions and recommended crop types for the given area.  
Output formats may include annotated maps, charts, or detailed reports, depending on user requirements.

## NON-FUNCTIONAL REQUIREMENTS:

### Performance:

The crop recommendation system should process high-resolution data, such as soil and environmental parameters, efficiently, with minimal latency, to provide timely and accurate crop suggestions. It should handle large datasets, including historical weather patterns, soil profiles, and market trends, without significant performance degradation..

### Accuracy:

The crop recommendation system should achieve high accuracy in identifying the most suitable crops based on soil, climate, and environmental factors while avoiding interference from irrelevant or non-impactful data such as inaccurate weather predictions or market anomal.

### Robustness:

The crop recommendation system should be robust against variations in environmental factors, such as soil conditions, weather patterns, and water availability. The system must adapt to various crops and farming techniques, ensuring its recommendations are accurate for different soil types, climate zones, and crop varieties

### Usability:

The user interface of the crop recommendation system should be intuitive and user-friendly, requiring minimal training for farmers to operate effectively.

The system’s outputs and visualizations should be clear and interpretable, ensuring that farmers can easily understand the crop recommendations.

## CONSTRAINTS:

### Hardware Limitations:

The performance of the crop recommendation system may be limited by the computational capacity and memory of the hardware available for running the system. This could affect the speed and efficiency of processing large datasets or real-time data inputs, such as weather patterns and soil conditions.

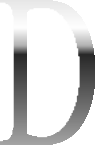
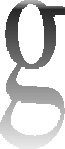
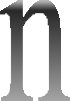
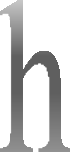
### Regulatory Compliance:

The crop recommendation system should comply with relevant data privacy regulations, particularly regarding the handling and storage of agricultural data, including sensitive information about farmers, their land, and crop details. Personal data collected from users must be protected in line with applicable data protection laws to ensure the privacy and security of individuals.

## DEPENDENCI

### Python Programming Language:

Software development and implementation shall be done using the Python programming language.



**CHAPTER-6**

SISTec/BTech/AD/2024/5/Minor Project\_I/03

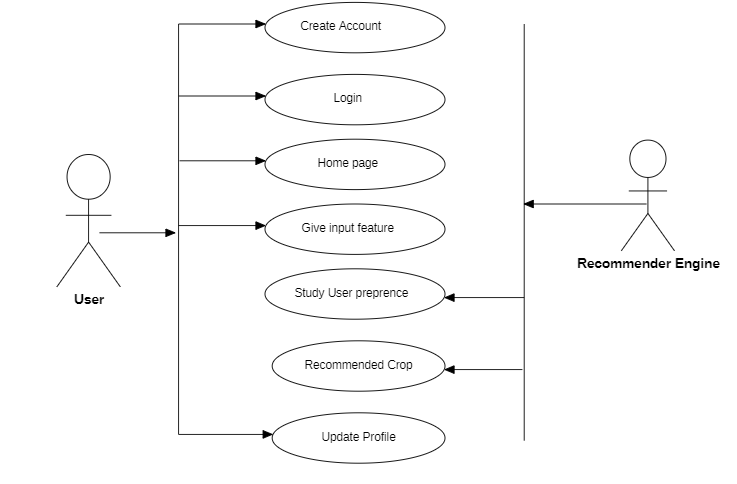
**SOFTWARE DESIGN**

## SOFTWARE DESIGN

Software design is a crucial phase in the development process where the system architecture and components are structured to meet the specified requirements. This chapter outlines the design aspects of the Crop Recommendation System.

### Use Case Diagram

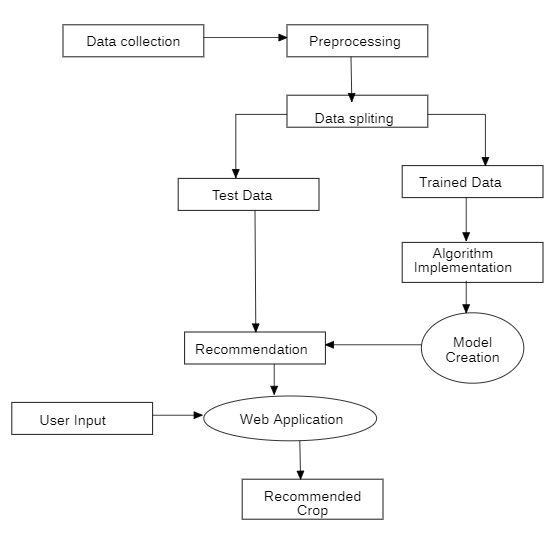
The Use Case Diagram illustrates the various interactions between users (actors) and the Crop Recommendation System. It identifies different use cases and describes how users interact with the system to achieve specific goals. In the context of the Crop Recommendation System, the actors may include:



### Figure 6.1 : Use Case Diagram

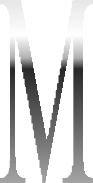
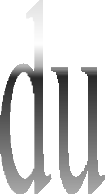
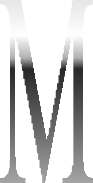
* 1. **Data Flow Diagram (DFD)**

The Data Flow Diagram (DFD) provides a visual representation of how data flows through the Crop Recommendation System. It illustrates the processes, data stores, and data flows within the system, highlighting the transformation of data at each stage. In the context of the Crop Recommendation System, the DFD may depict the flow of data from the input sources, such as soil testing results, climate data, and farmer inputs, to the preprocessing module, crop suitability analysis, recommendation generation, and output visualization or report generation.



**Figure 6.2 : Data Flow Diagram**

SISTec/BTech/AD/2024/5/Minor Project\_I/03



**CHAPTER-7**

**MACHINE LEARNING MODULE**

## DATASET DESCRIPTION

The dataset for the Crop Recommendation System project was sourced from a Kaggle website. Soil and environmental data were collected from diverse agricultural regions, representing various soil types, climatic conditions, and farming practices. This data was gathered under standardized procedures to ensure accuracy and consistency, which is essential for reliable model training and prediction tasks.

The dataset includes soil samples with varying nutrient levels, pH values, and moisture content, along with data on climatic factors like temperature, rainfall, and humidity. It also captures information about diverse crop types, including commonly grown staples and niche crops, to account for regional variations.This diversity allows the system to generalize effectively across different agricultural regions and conditions, ensuring its practical applicability for farmers across various geographies and resource levels.

## PRE-PROCESSING STEPS

To ensure the dataset's quality and suitability for training, several pre-processing steps were applied:

**Data Cleaning:**  Handle missing values by using techniques like imputation (mean, median, or mode) or removing rows/columns with excessive missing data.Remove duplicate records to avoid redundancy.

**Outlier Detection and Handling:** Identify outliers using methods like Z-score, IQR, or visualizations (box plots). Decide whether to remove or transform outliers based on their relevance.

**Encoding**: Convert categorical variables into numerical forms using methods like One-Hot Encoding, Label Encoding, or Frequency Encoding.

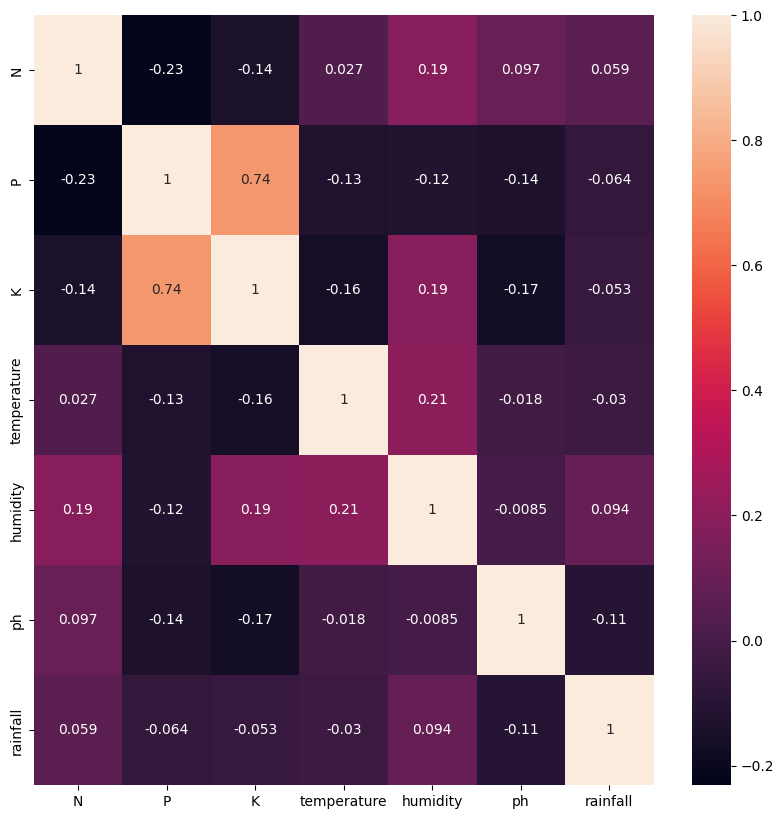
**Scaling/Normalization**: Adjust data to a common scale using techniques like Min-Max Scaling, Standardization, or Robust Scaling.

**Data Splitting:** Split the dataset into training, validation, and test sets to ensure proper evaluation of the model.

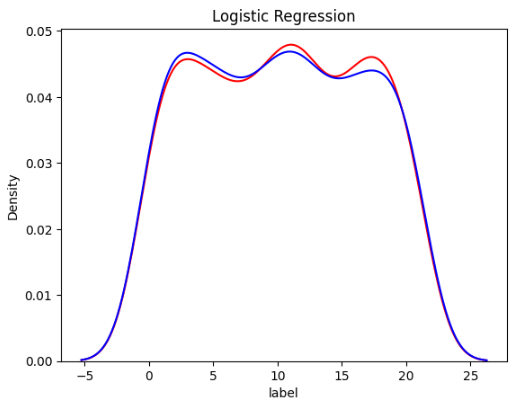
## DATA VISUALIZATION



### Figure 7.1 : Pair plot



**Figure 7.2 : Confusion Matrix**



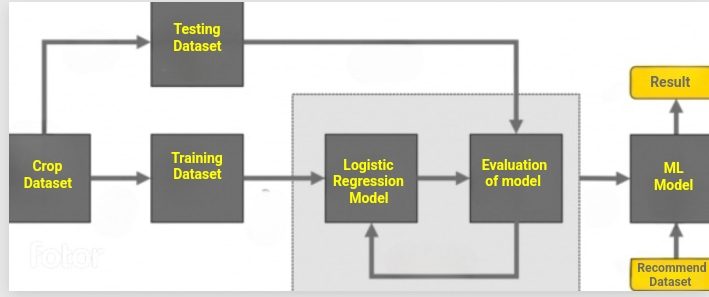
**Figure 7.3 : Prediction Curve**

## ML MODEL DESCRIPTION

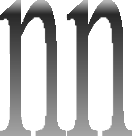
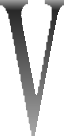
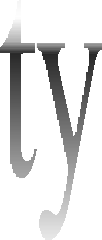
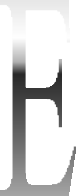
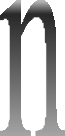
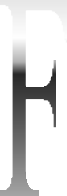
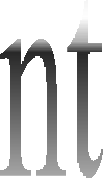
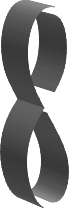
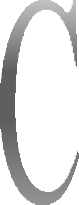
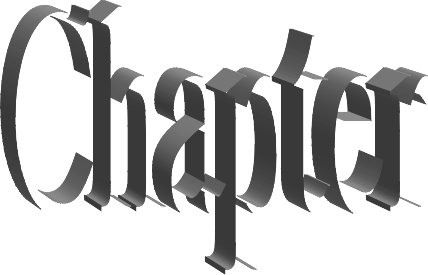
For this project, we employed Logistic Regression, a widely-used statistical model known for its effectiveness in multi-class classification tasks. Logistic Regression provides a robust balance between interpretability and accuracy, making it suitable for applications requiring categorical predictions. Its ability to handle multiple classes allows it to classify distinct categories effectively, which is essential for our purpose of analyzing and categorizing crop features.

The choice of Logistic Regression as the classification model was based on its adaptability to various multi-class problems and its efficiency in modeling complex relationships between input features and target classes. By training the model on our labeled clinical crop dataset, we aimed to enhance its ability to detect subtle features, such as minor changes in soil quality, pest infestations, or crop health, even under diverse field conditions. This enables precise crop recommendations tailored to the specific needs of farmers.

## MODEL WORKFLOW

****

### Figure 7.4 Model Workflow



**CHAPTER-8**

SISTec/BTech/AD/2024/5/Minor Project\_I/03

**FRONT-END CONNECTIVITY**

Connecting a machine learning (ML) model to a frontend involves creating a pipeline that allows users to interact with the model through a web interface or application. Here’s an overview of the process:

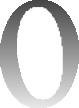
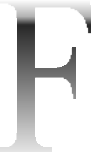
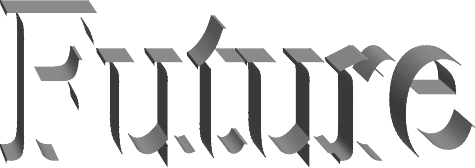
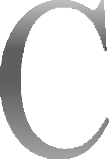
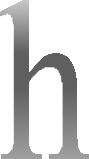
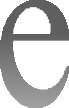
## STEPS OF FRONT-END CONNECTIVITY

1. **HTML Structure & Layout:** The HTML defines the layout for the Crop Recommendation system, including sections for navigation, an user input feature, and an about section. CSS (in "styles.css") is used to style the layout.
2. **User Input:** A user can press recommend button to allow users to fill the form.
3. **Predict Button & Request Handling:** After filling form, a "Recommend Crop" button appears. When clicked, the Recommendation function is triggered, which prepares the crop data in a FormData object.
4. **Data Transfer to Backend (Django Server):** The crop data is sent to the backend Django server (http://localhost:8000/detect) through a fetch API POST request for processing.
5. **Displaying Results:** Once the server returns a response containing the Recommend best Crop for cultivation.
6. **Error Handling:** If there’s an issue with the recommendation request or response, an alert is shown, informing the user to try again.

## 8.2 ****Backend Setup for the ML Model****

1. **Model Deployment**: The trained ML model in a compatible format (e.g, .pkl for Scikit-learn). Use frameworks Django, to create an API for serving predictions.

SISTec/BTech/AD/2024/5/Minor Project\_I/03



**CHAPTER-9 FUTURE WORK**

## INTRODUCTION

In this chapter, we explore potential avenues for future enhancements and developments in the Crop Recommendation System (CRS). By identifying areas for improvement and novel research directions, we aim to advance the capabilities and effectiveness of the CRS in accurately recommending suitable crops. Future enhancements may include integrating additional data sources, such as satellite imagery and IoT-based sensors, enabling real-time data processing, and employing advanced machine learning models to improve recommendation accuracy. Additionally, expanding the system’s functionality to include pest and disease prediction, as well as climate change adaptability, represents a valuable opportunity for broader agricultural applications.

## INTEGRATION OF ADDITIONAL DATA MODALITIES

One direction for future work involves the integration of additional data sources beyond the soil, weather, and environmental factors currently used in the Crop Recommendation System. Incorporating advanced technologies such as remote sensing, multispectral satellite imaging, or drone-based monitoring could provide complementary data, enhancing the precision of crop recommendations and improving decision-making under varying soil types, climatic conditions, and cropping patterns.

## REAL-TIME PROCESSING AND ANALYSIS

Enhancing the Crop Recommendation System (CRS) to enable real-time processing and analysis represents a promising area for development. Implementing real-time data analysis capabilities would allow immediate recommendations and insights, enabling farmers to obtain instant feedback for applications such as crop selection, soil health analysis, and weather-based planting decisions.

## ADVANCED MACHINE LEARNING MODELS

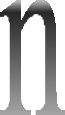
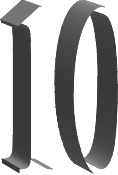
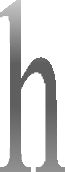
Future iterations of the crop recommendation system could benefit from advanced machine learning models and techniques. Exploring architectures such as transformer-based models or generative adversarial networks (GANs) could enhance the accuracy of crop predictions, particularly in challenging conditions such as overlapping environmental factors or.

inconsistent data inputs (e.g., varying soil quality or weather patterns). Fine-tuning these models with larger and more diverse datasets could improve the system's robustness and adaptability to a wider range of agricultural scenarios

## CONCLUSION

The future work outlined in this chapter highlights the potential for continued refinement and expansion of the Crop Recommendation System. By leveraging advancements in data collection technologies, machine learning, and additional environmental and crop-specific features, researchers can enhance the system’s accuracy, adaptability, and practical usability. This progress would enable more precise recommendations tailored to diverse agricultural conditions, thereby broadening its impact and effectiveness in supporting sustainable farming practices.

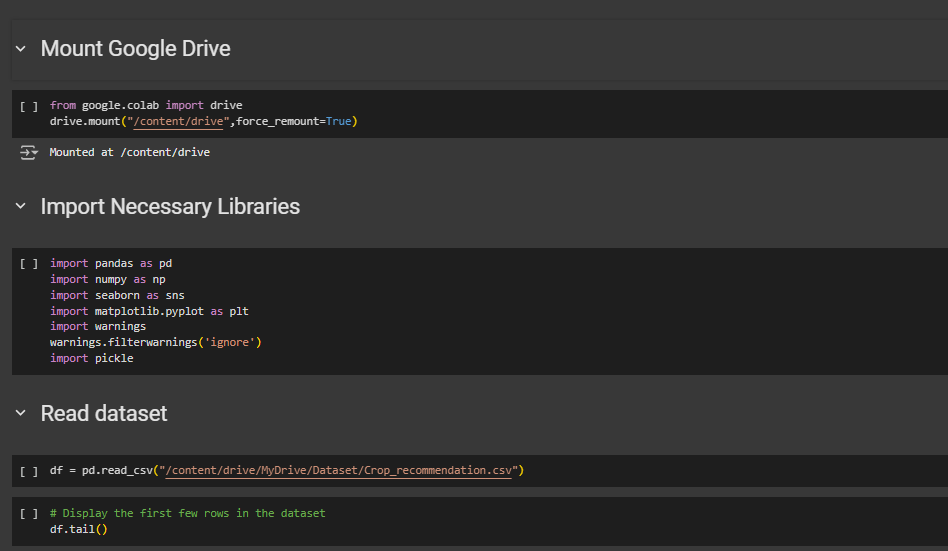
SISTec/BTech/AD/2024/5/Minor Project\_I/03

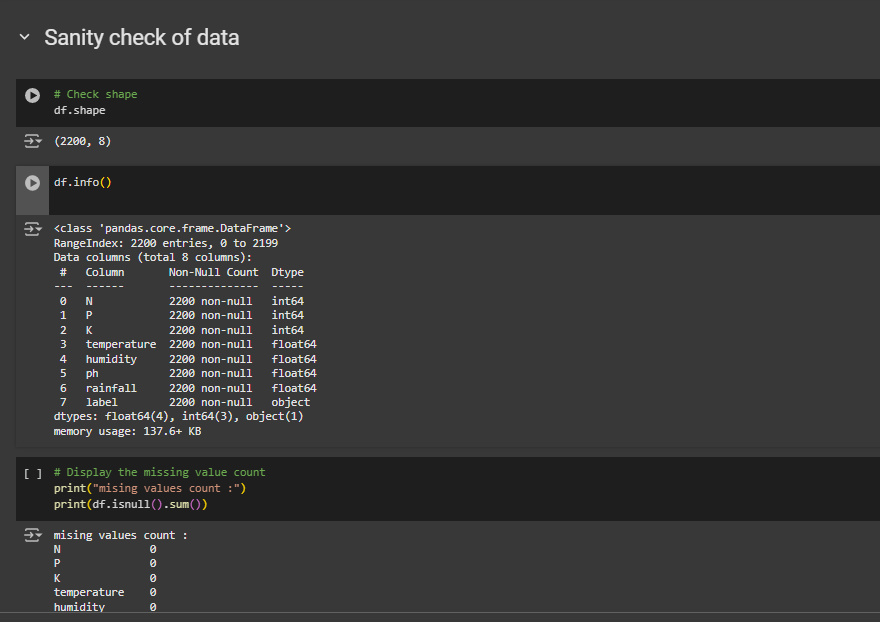


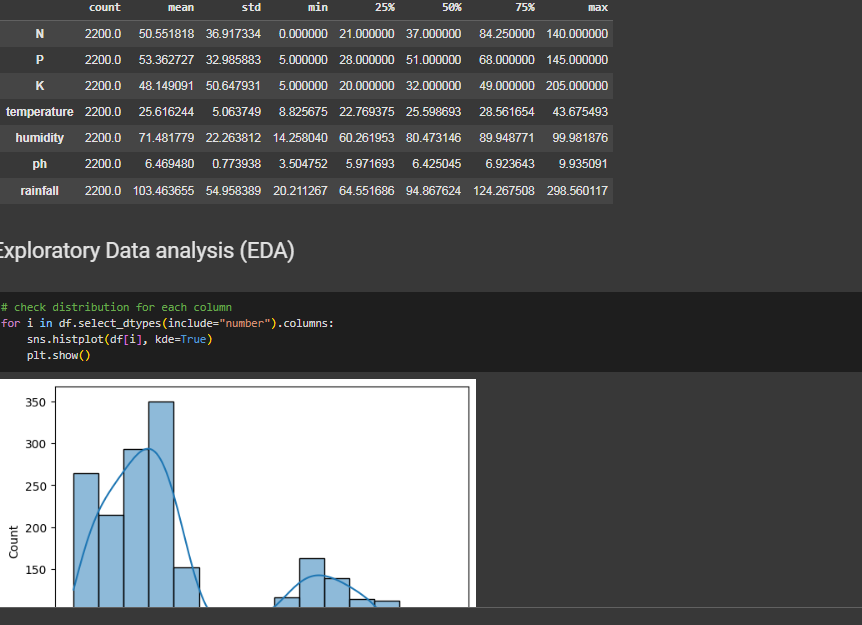
**CHAPTER-10**

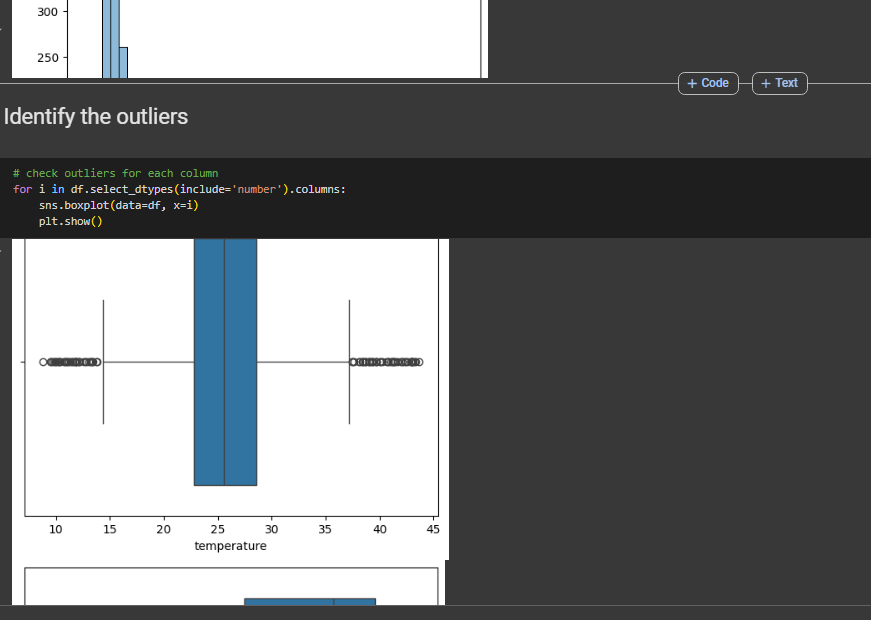
**Coding**

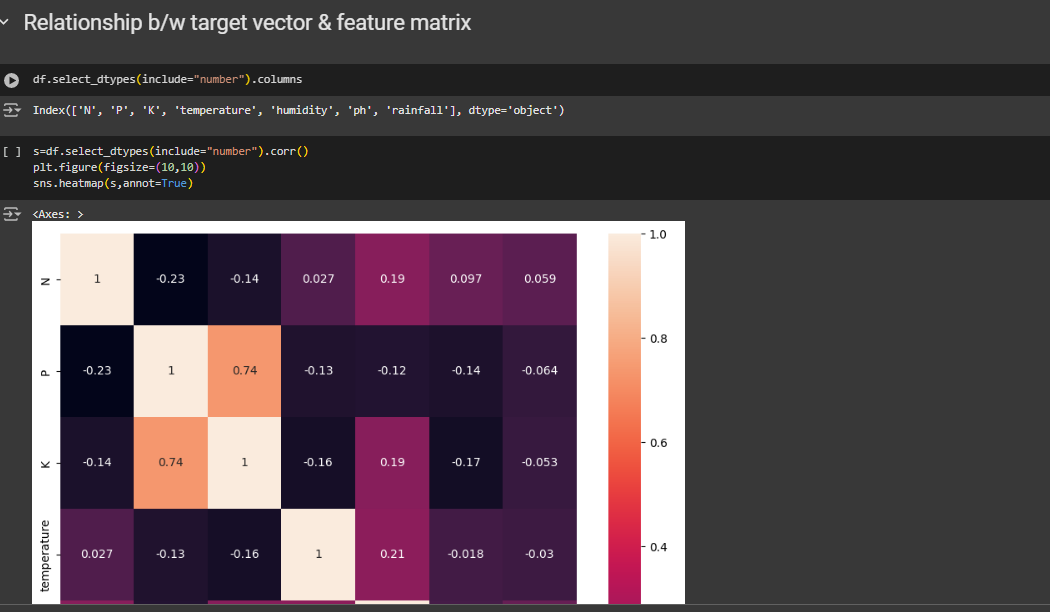
**ML Model**

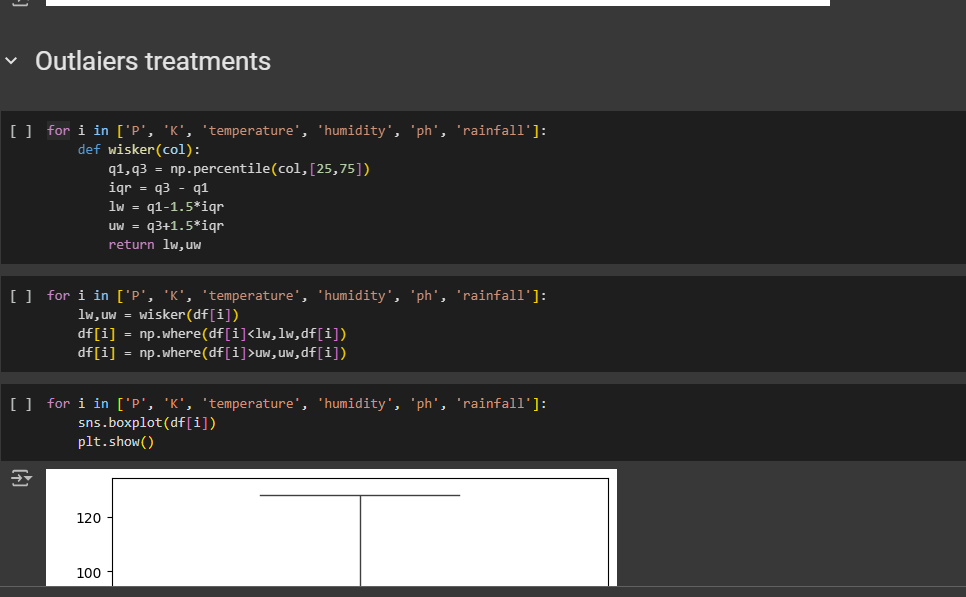
****

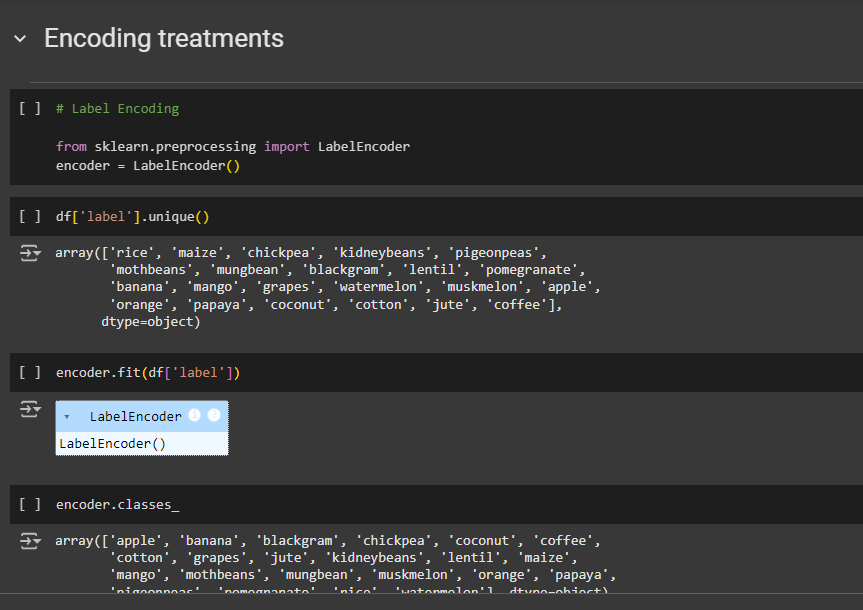
****

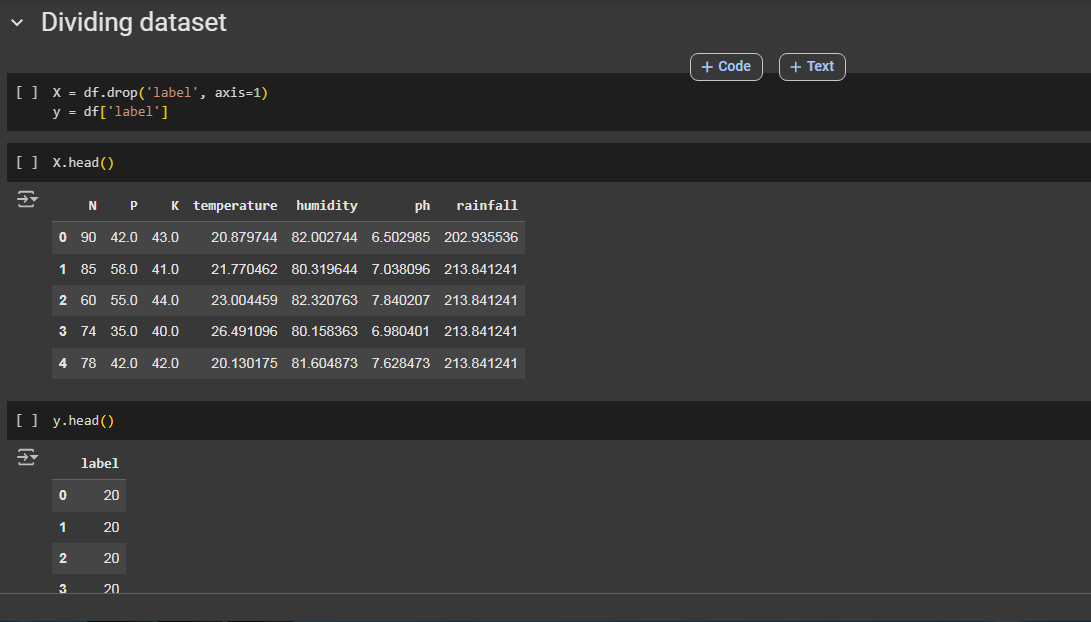
****

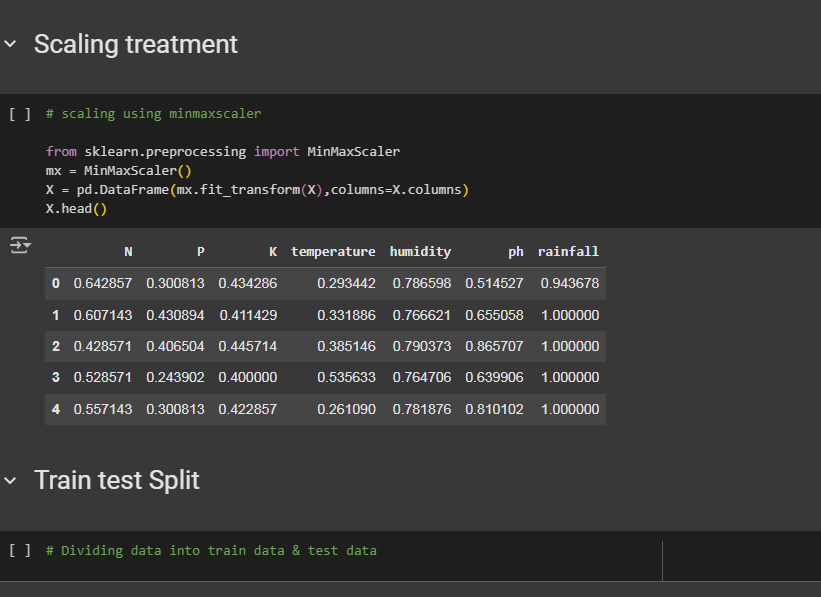
****

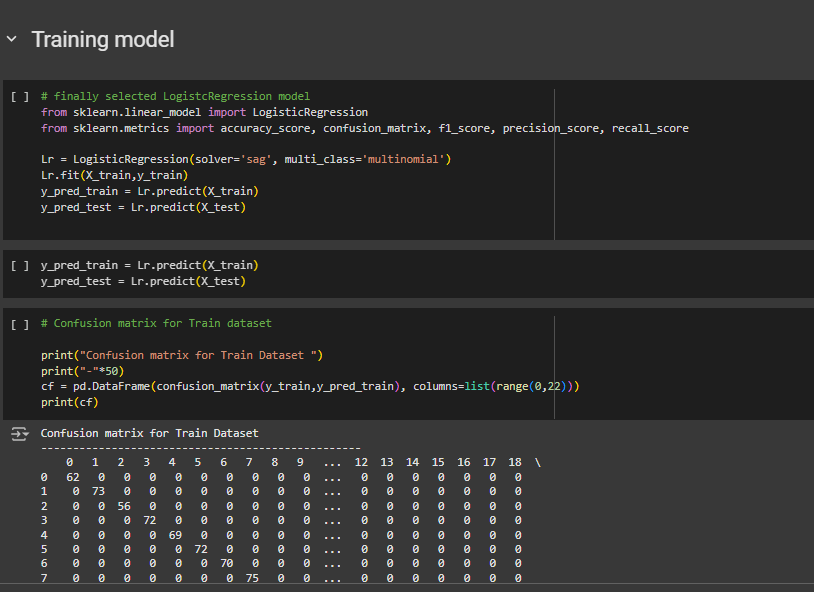
****

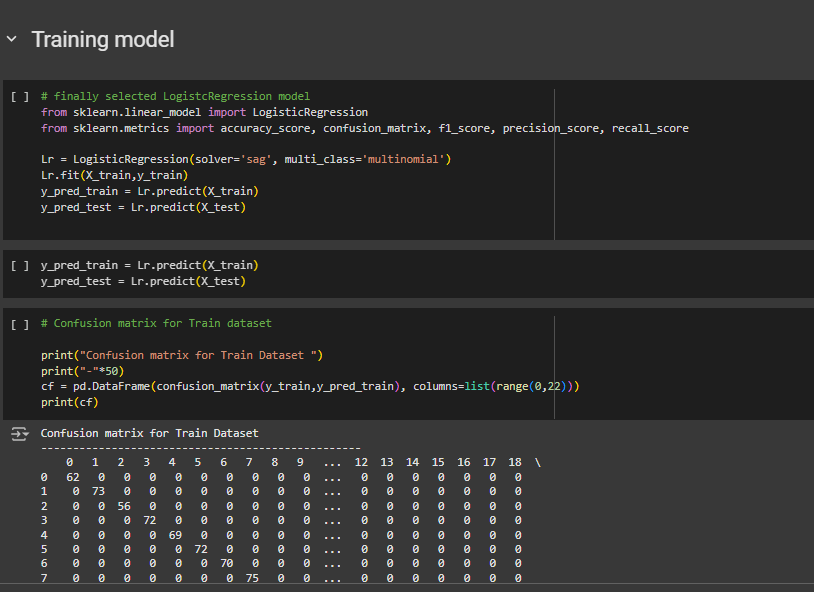
****

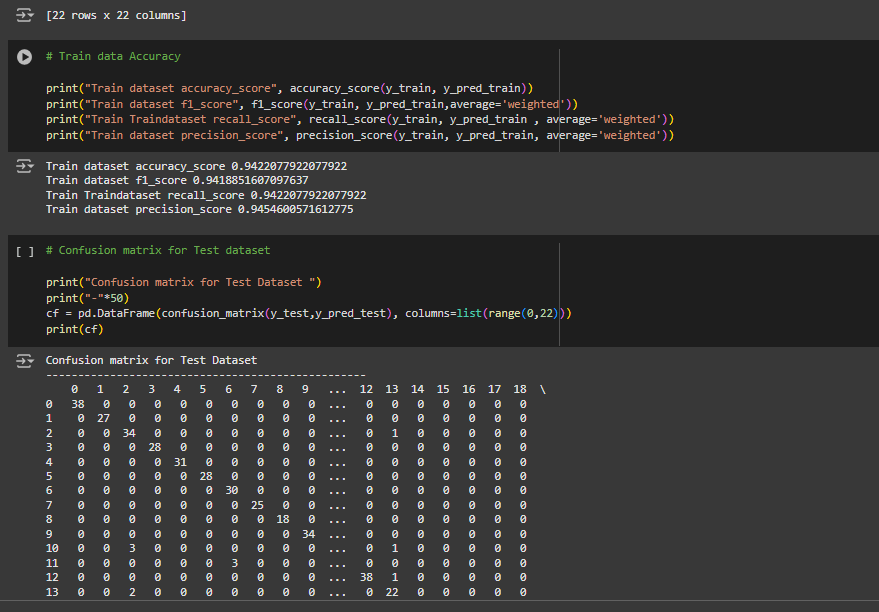
****

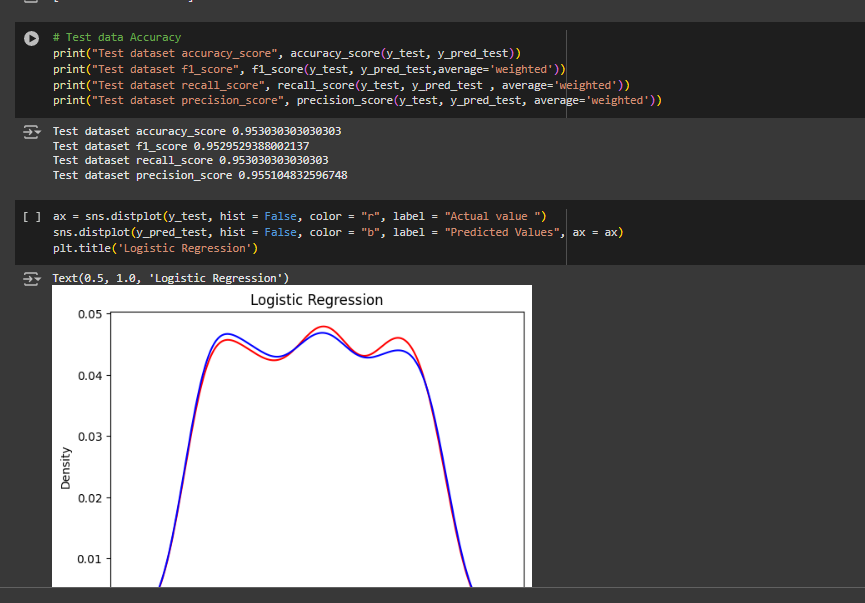


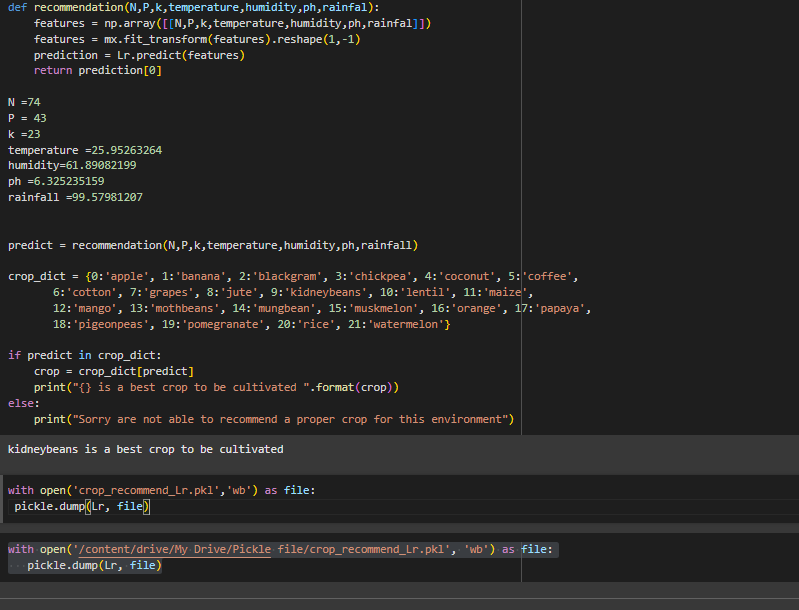


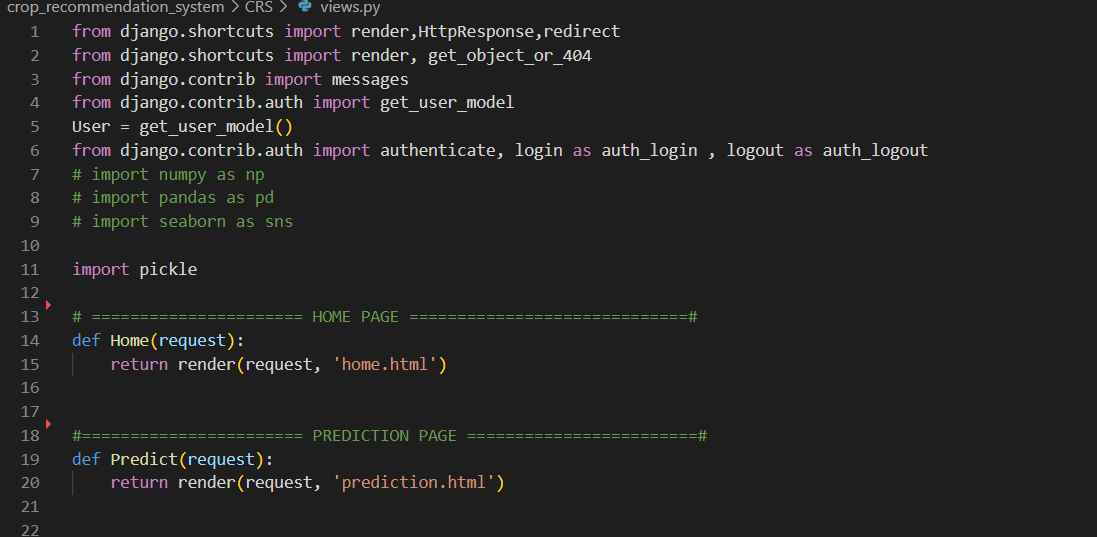


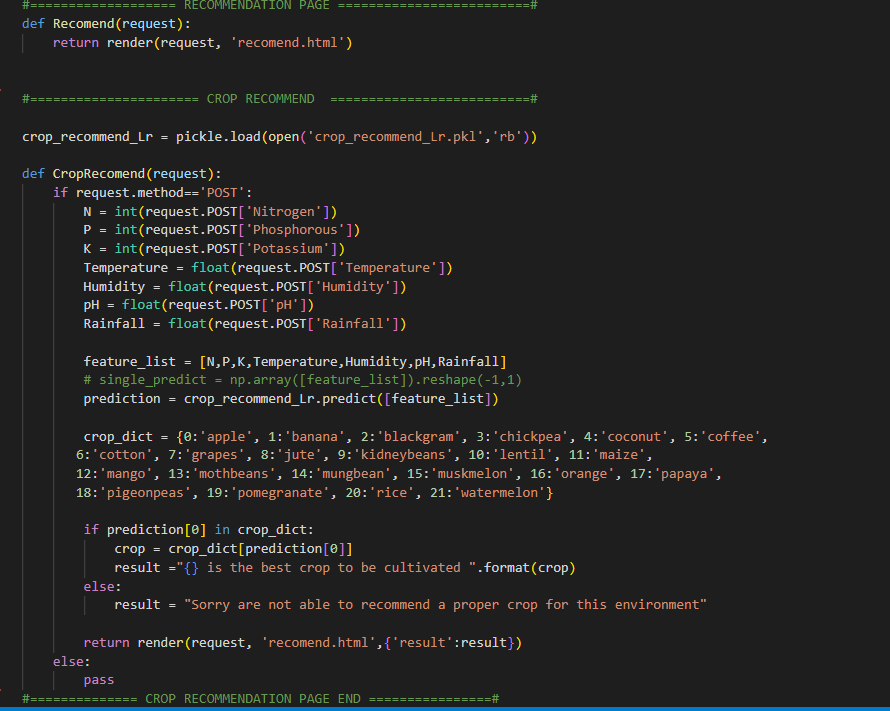


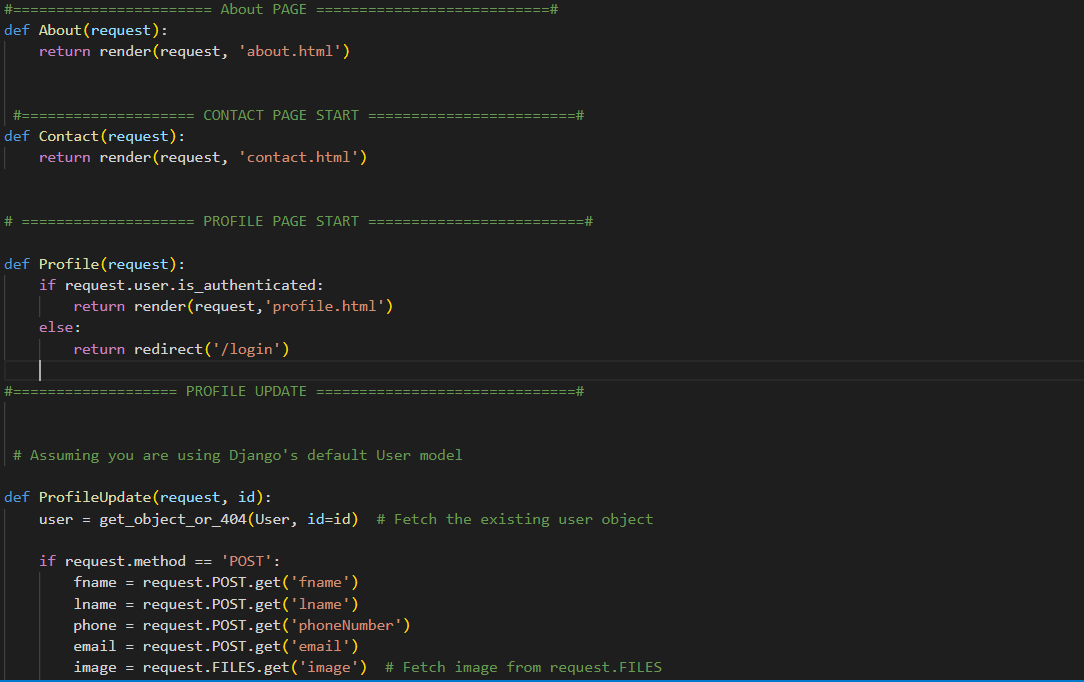


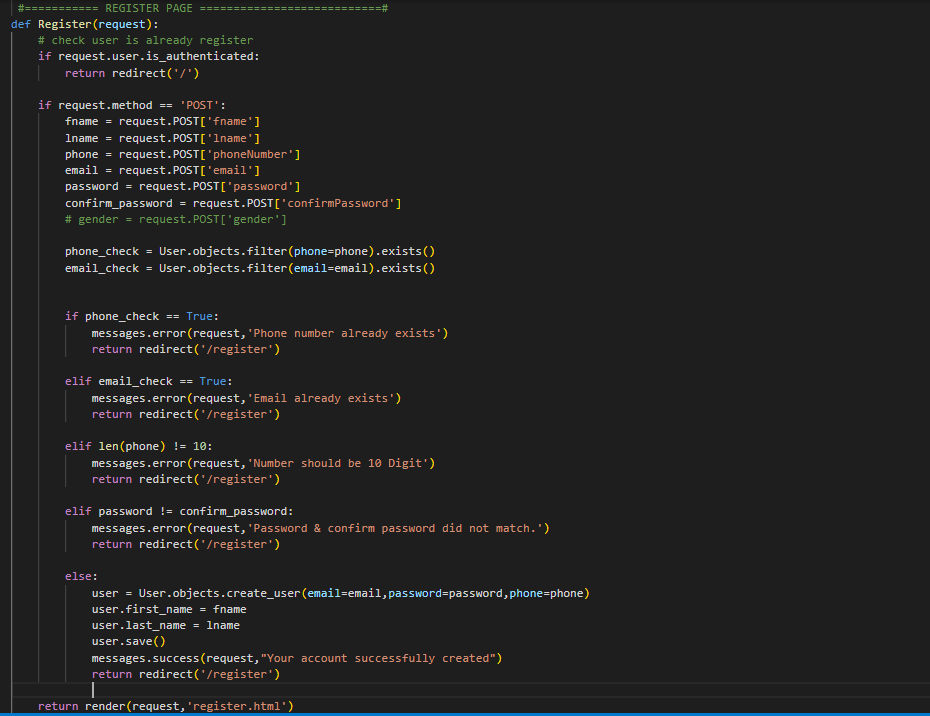


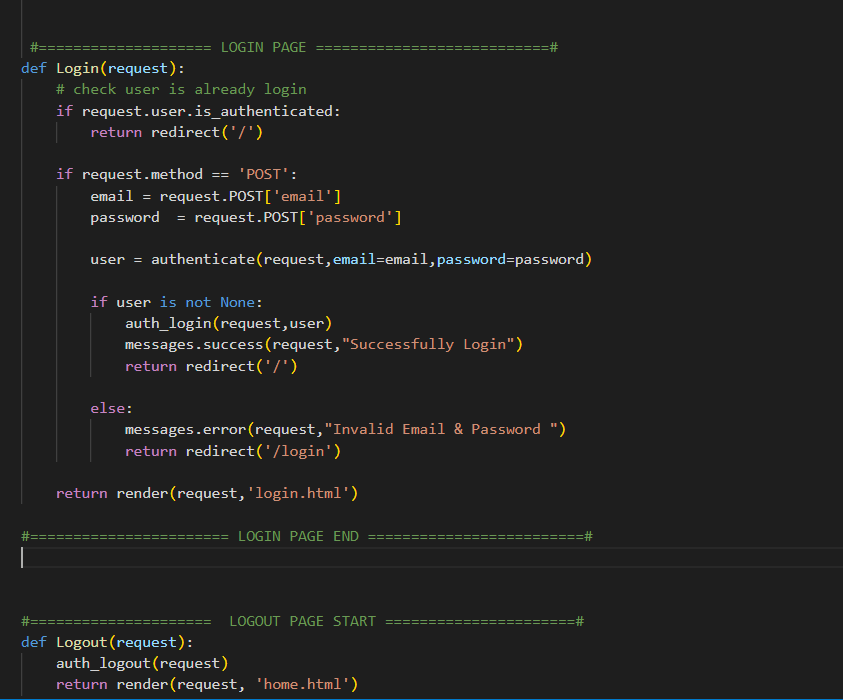


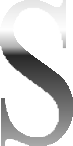
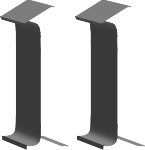
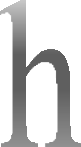








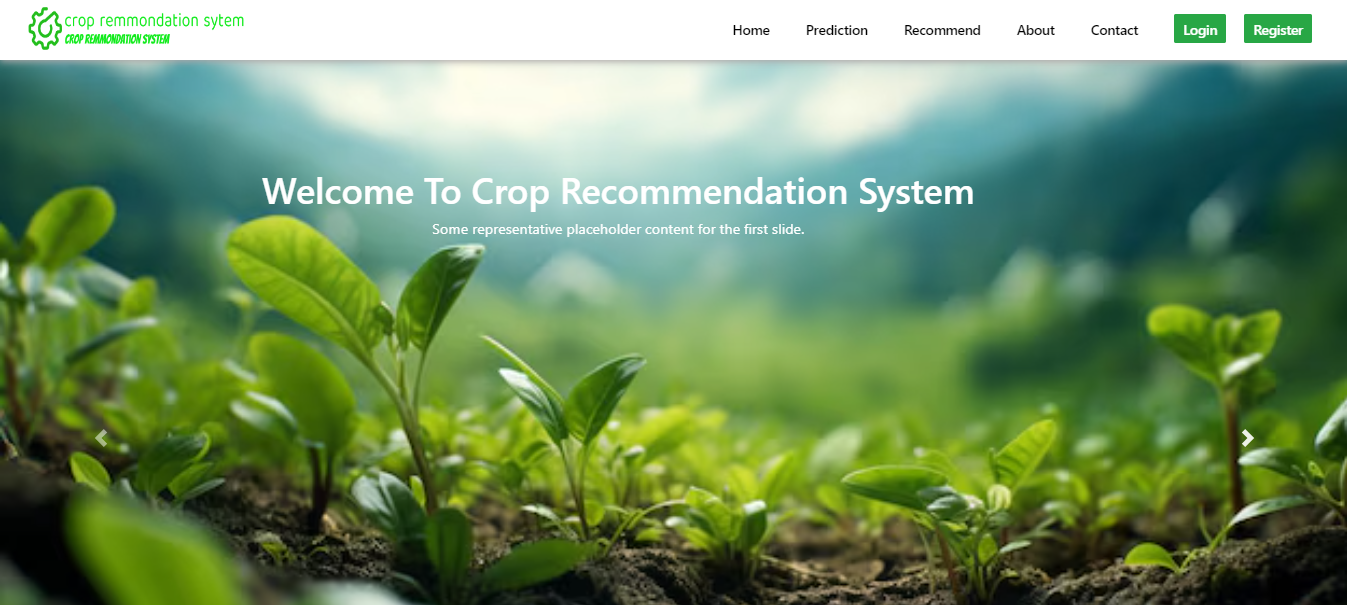




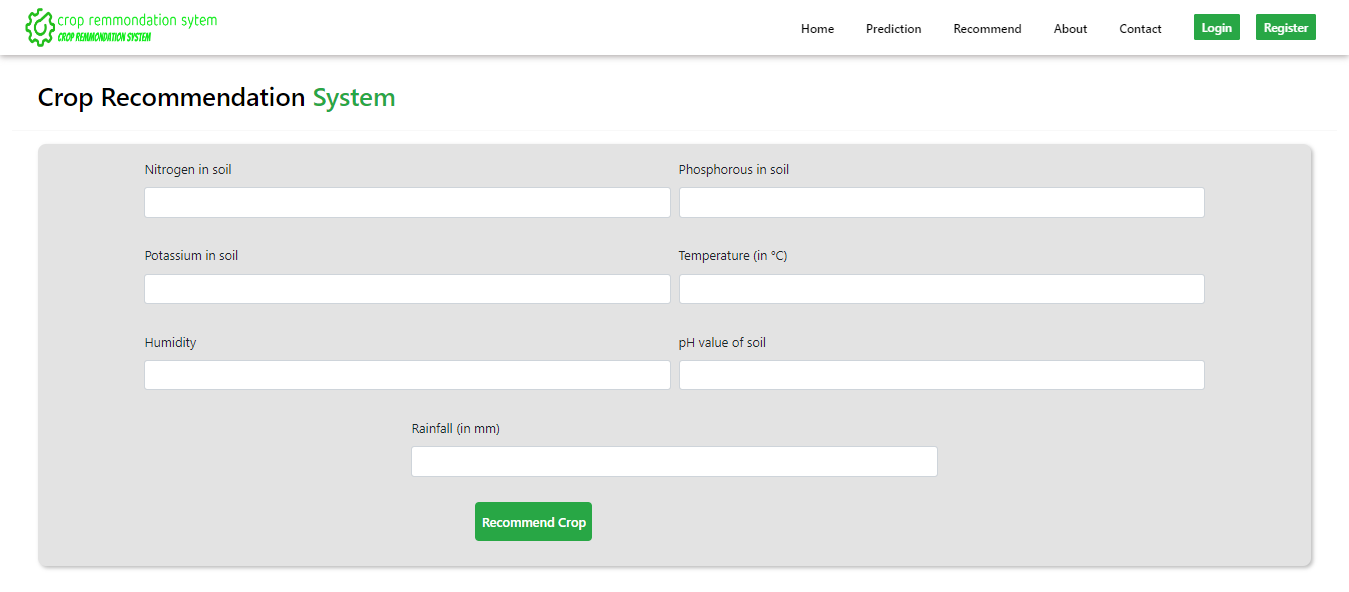
**CHAPTER-11**

**Output Screen**

## HOME PAGE



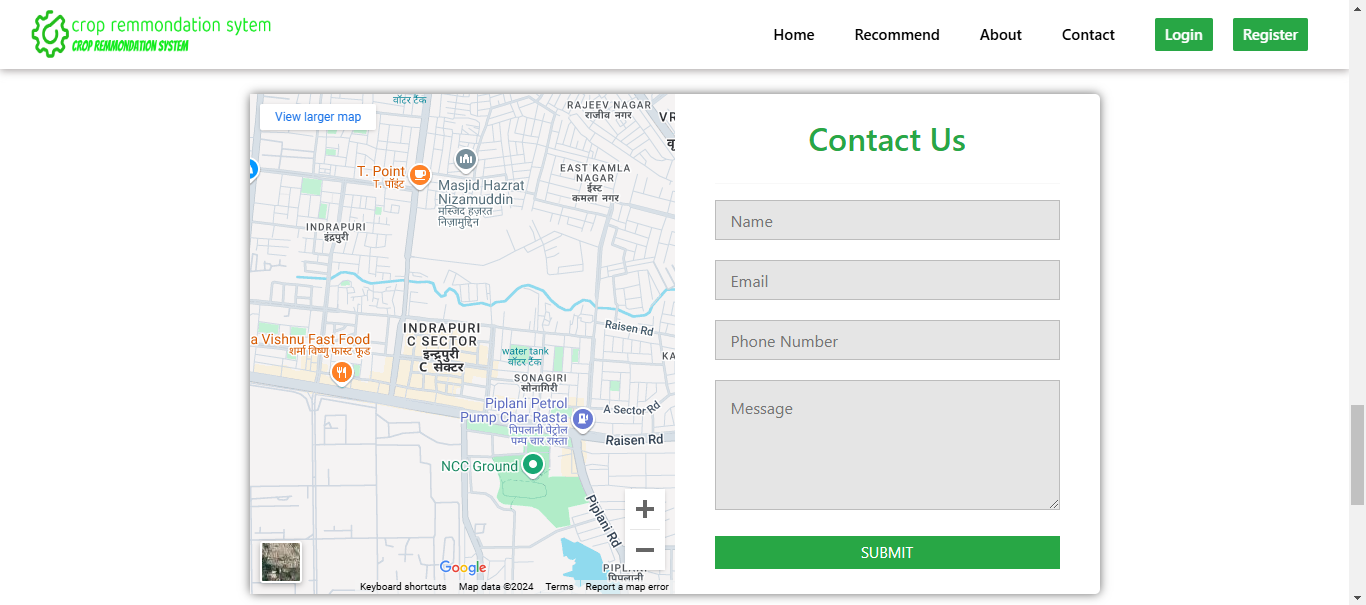
1. **RECOMMENDATION PAGE**



## OUR SERVICES



1. **CONTACT PAGE**



**REFERENCES**

1. Joseph Berkson, ―Logistic Regression, 2019 IEEE Transactions on Instrumentation and Measurement.
2. <https://en.wikipedia.org/wiki/Crop>
3. <https://www.geeksforgeeks.org/machine-learning/>
4. <https://docs.djangoproject.com/en/5.1/>
5. <https://scikitlearn.org/dev/modules/generated/sklearn.linear_model.LogisticRegression.html>
6. <https://chat.openai.com/>
7. https://www.kaggle.com/datasets
8. <https://github.com/>

**APPENDIX-1 GLOSSARY OF TERMS**

|  |  |
| --- | --- |
| **B** |  |
| **Bias in Data** | Unintended patterns in the data that can affect the model's predictions, often requiring mitigation during preprocessing to avoid skewed  results. |
| **D** |  |
| **Data Collection** | The process of gathering relevant data for training a crop recommendation system using an ML model involves collecting diverse and high-quality datasets related to agricultural practices. |
| **Data Preprocessing** | Cleaning and preparing the raw data to ensure its quality and suitability for the machine learning model’s training. |
| **Machine Learning** | We used Logistic regression model |
| **F** |  |
| **Feature Engineering** | Feature Engineering: The process of selecting and creating relevant features from the dataset to improve the model's predictive  accuracy. |
| **Feature Extraction** | The technique used in data cleaning to extract relevant information from raw data for use in machine learning models. |
| **H** |  |
| **Hosting Environment** | The platform or cloud service (such as AWS or  Google Cloud) where the trained model and the |

|  |  |
| --- | --- |
|  | Crop recommendation system will be deployed |
| **M** |  |
| **Model Evaluation** | The process of assessing the performance and accuracy of the machine learning model using various metrics, such as precision, recall, and  F1-score, specifically in crop recommendation system. |
| **Monitoring and Maintenance** | A plan for continuous monitoring and upkeep of the model, ensuring the system remains accurate and efficient over time, particularly when  handling real-world data. |
| **S** |  |
| **Scalability** | The system's ability to process a growing number of crop recommend data and adapt to increasing amounts of data without significant  performance degradation. |
| **U** |  |
| **User Interface** | The graphical interface that allows users to interact with the crop recommendation system, such as fill use data for analysis or  viewing recommended crop results. |
| **V** |  |
| **Version Control** | The practice of tracking changes made to code, models, and other project files, often using tools like Git, to ensure proper management of  project development and updates. |

**Appendix 2 – PROJECT SUMMARY**

# About Project

|  |  |
| --- | --- |
| **Title of the project** | Design and Development of Crop Recommendation System  using Machine Learning |
| **Semester** | 5 |
| **Members** | 2 |
| **Team Leader** | Aman Sahu |
| **Describe role of every member in the project** | **Aman Sahu** - is responsible for training the machine learning model for the Crop Recommendation System, focusing on developing and fine-tunning  **Shyam Sahu** - is responsible for sourcing and preparing the dataset required for training the Crop Recommendation System, including gathering data of crop Recommendation system. |
| **What is the motivation for selecting this project?** | The motivation for selecting the **Crop Recommendation System** project stems from the growing need for advanced technologies in the field of **agriculture**. Accurate and automated analysis of soil, climate, and crop-related data can significantly enhance farming practices, improve personalized crop selection for farmers, and provide a data-driven, non-invasive approach to optimize agricultural productivity and ensure sustainable farming practices. |
| **Project Type**  **(Desktop Application, Web Application, Mobile App, Web)** | Web Application |

***Tools &Technologies***

|  |  |
| --- | --- |
| **Programming language used** | Python |
| **Interpreter used** | Python (3 or above ) |

|  |  |
| --- | --- |
| **IDE used** | Jupyter Notebook 7.1.2 |

# Project Requirements

|  |  |
| --- | --- |
| **MVC architecture followed**  **(Yes / No)** | Yes |
| **If yes, write the name of MVC architecture followed**  **(MVC-1, MVC-2)** | MVC-2 |
| **Design Pattern used**  **(Yes / No)** | No |
| **If yes, write the name of Design Pattern used** |  |
| **Interface type**  **(CLI / GUI)** | GUI |
| **No. of Actors** | 1 |
| **Name of Actors** | User |

***Testing***

|  |  |
| --- | --- |
| **Which testing is performed?**  **(Manual or Automation)** | Automation |
| **Is Beta testing done for this project?** | NO |

# Write project narrative covering above mentioned points

The motivation for selecting the Crop Recommendation System project arises from the increasing demand for advanced technologies in agriculture to improve productivity, sustainability, and decision-making. Crop selection is a critical aspect of farming, and accurate, automated recommendations offer great potential for enhancing yields, resource optimization, and overall agricultural efficiency. Traditional methods of selecting crops often rely on intuition or manual assessments, which can be time-consuming, subjective, and limited in accuracy. This project aims to leverage machine learning and data analysis techniques to develop an innovative system for real-time crop recommendations based on soil, weather, and environmental conditions. By doing so, it will enable more precise and objective decision-making for farmers, reducing the risks of crop failure and maximizing profitability. The system could be particularly beneficial in agricultural settings, assisting farmers, agronomists, and policymakers to implement data-driven strategies for better crop management and sustainable farming practices.

|  |  |
| --- | --- |
| ***Aman Sahu*** | ***Shyam Sahu*** |
| 0187AS221007 | 0187AS221059 |