A Project Report on

"Intelligent Crop Recommendation System Using Machine Learning"

Submitted by:

Miss. TEJAL D. SURVASE (PRN:2019032500207344)
Miss. SIDDHI S. VELAPURE (PRN:2019032500207197)
Miss. MAYURI S. SWAMI (PRN:2019032500209173)
Miss. RATAN G. KORE (PRN:2019032500210192)
Miss. PRAPTI P. ATKALE (PRN:2019032500207441)

UNDER THE GUIDANCE OF

Mr. P. D. Mane

In partial fulfillment for the award of the

degree of

BACHELOR

OF ENGINEERING IN

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

at

SHRI VITHAL EDUCATION and RESEARCH INSTITUTES's, COLLEGE OF ENGINEERING,PANDHARPUR

AFFILIATED TO PUNYASHLOK AHILYADEVI HOLKAR SOLAPUR UNIVERSITY,SOLAPUR

2022-2023

A Project Report on

"Intelligent Crop Recommendation System Using Machine Learning"

Submitted by:

Miss. TEJAL D. SURVASE (PRN:2019032500207344)
Miss. SIDDHI S. VELAPURE (PRN:2019032500207197)
Miss. MAYURI S. SWAMI (PRN:2019032500209173)
Miss. RATAN G. KORE (PRN:2019032500210192)
Miss. PRAPTI P. ATKALE (PRN:2019032500207441)

UNDER THE GUIDANCE OF

Mr. P. D. Mane

In partial fulfillment for the award of the

degree of

BACHELOR

OF ENGINEERING IN
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
at

SHRI VITHAL EDUCATION and RESEARCH INSTITUTES's, COLLEGE OF ENGINEERING,PANDHARPUR

AFFILIATED TO PUNYASHLOK AHILYADEVI HOLKAR SOLAPUR UNIVERSITY,SOLAPUR

2022-2023

SVERI’s COLLEGE OF ENGINEERING, PANDHARPUR

CERTIFICATE

This is to certify that the project report entitled **“Intelligent Crop Recommendation System Using Machine Learning”** is submitted for partial fulfillment of Bachelor Of Technology as per requirement of Punyashlok Ahilyadevi Holkar Solapur University, Solapur for the academic year 2022-2023.

Mr. P. D. Mane
(Project Guide)

Mr. P. D. Mane
(Project Coordinator)

Dr. S. P. Pawar
(HOD, CSE)

Dr. B. P. RONGE
(PRINCIPAL)

EXTERNAL EXAMINAR

Acknowledgement

We are pleased to acknowledge Dr. S. P. Pawar(HOD CSE) sir for his valuable guidance during the course of this project work. We extend our sincere thanks to Mr. P. D. Mane who continuously helped us throughout the project and without his guidance, this project would have been an uphill task. We are also grateful to other members of the CSE faculty members and technical staff who cooperated with us regarding some issues.

Last but not the least, Ms. T. D. Dhumal madam supervisor of Project Lab as the case may be for project sessions, are also cooperated with us nicely for the smooth development of this project. I would also like to thank my parents and friends who helped me a lot in executing this project within the limited time frame.

Signature

Miss. Survase T. D. Sign. ....
Miss. Velapure S. S. Sign. ....
Miss. Swami M. S. Sign ....
Miss. Kore R. G. Sign. ....
Miss. Atkale P. P. Sign ....

SYNOPSIS

1. Synopsis of Project Work

(a) Relevance:

An intelligent crop recommendation system is highly relevant in the context of agriculture due to several key factors: Changing climate condition. Climate change is affecting agricultural systems worldwide, with shifts in temperature, rainfall patterns, and extreme weather events. An intelligent crop recommendation system can analyze historical climate data and predict future climate trends to suggest crop varieties that are more resilient and adaptable to changing conditions.

1.Soil variability: Soils differ in their fertility, composition, and water-holding capacity across regions. A crop recommendation system can take into account soil characteristics, such as pH levels, nutrient content, and organic matter, to recommend crops that are suitable for specific soil types. This helps optimize crop growth and minimize soil degradation.

2.Market demand and profitability: The agricultural market is dynamic, with varying demand for different crops and fluctuating prices. An intelligent crop recommendation system can consider market trends, consumer preferences, and crop profitability to guide farmers in selecting crops that have higher market demand and better economic returns.

3.Sustainable practices: There is an increasing focus on sustainable agriculture practices to conserve resources and minimize environmental impact. A crop recommendation system can suggest crops that require fewer pesticides, fertilizers, or irrigation, thereby promoting sustainable farming methods and reducing the use of agrochemicals.

4.Precision agriculture: Precision agriculture involves the use of technologies like remote sensing, drones, and sensors to collect data on crop health, soil moisture, and nutrient levels. An intelligent crop recommendation system can integrate these data sources to provide real-time recommendations for precise fertilizer application, irrigation scheduling, and pest management, leading to improved efficiency and reduced resource waste.

5.Farmer empowerment: An intelligent crop recommendation system empowers farmers by providing them with valuable information and insights. It assists farmers in making data-driven decisions based on scientific analysis, reducing guesswork, and improving overall productivity and profitability.

6.Scaling up agricultural operations: As agriculture expands to meet the growing global population's food demand, there is a need for efficient and sustainable farming methods. An intelligent crop recommendation system can help scale up agricultural operations by optimizing crop choices, resource allocation, and management practices, ensuring that agriculture can meet the increasing food requirements without compromising environmental sustainability.

In summary, an intelligent crop recommendation system is relevant because it addresses the challenges and complexities faced by modern agriculture, including climate change, soil variability, market dynamics, sustainability goals, precision agriculture, and the need for farmer empowerment. By leveraging technology and data-driven insights, such a system can improve agricultural decision-making, increase productivity, and contribute to a more sustainable and resilient food production system.

(b)Present theories and practices :

1.IBM AgroPad: IBM AgroPad is a technology developed by IBM Research that combines a handheld device with machine learning algorithms to analyze soil and water samples. It provides smallholder farmers with recommendations on crop selection and optimal water and fertilizer usage.

2.Taranis: Taranis is an agtech company specializing in aerial imagery and AI-driven analytics. They leverage high-resolution imagery, weather data, and machine learning algorithms to identify crop stress, pests, and diseases. Taranis provides farmers with recommendations for targeted treatments and crop management decisions.

3.CropX: CropX is a soil sensing and analytics company that offers a soil monitoring system to optimize irrigation and fertilizer management. Their platform combines soil sensors, satellite imagery, and machine learning algorithms to provide farmers with real-time insights and recommendations for irrigation scheduling and nutrient application.

4.SST Software (Sirrus): SST Software provides the Sirrus platform, which combines data analytics, remote sensing, and mapping tools to assist farmers with crop management decisions. It offers recommendations for seed selection, planting rates, variable rate applications, and field scouting.

(c) Proposed work :

i. Scope of the project:

Intelligent crop recommendation systems involves the integration of emerging technologies, personalized recommendations, climate change adaptation, advanced data analytics, precision agriculture practices, and collaborative platforms. These advancements will enhance the efficiency, productivity, and sustainability of agricultural practices worldwide.

ii. Objectives :

The main objective of the system is to enhance agricultural productivity, sustainability, and profitability while addressing environmental concerns and enabling farmers to make informed decisions for their farming operations.

1.Improve Crop Productivity: The primary objective of an intelligent crop recommendation system is to enhance crop productivity and optimize yield potential. By providing accurate recommendations on suitable crop varieties, optimal planting dates, and appropriate agronomic practices, the system aims to maximize crop yields and overall farm productivity.

2.Enhance Crop Quality: Intelligent crop recommendation systems also focus on improving the quality of harvested crops. By considering factors such as market demand, consumer preferences, and crop traits, the system suggests suitable varieties and cultivation practices that result in higher-quality produce, thereby meeting market standards and increasing profitability for farmers.

3.Enable Climate Change Adaptation: With changing climate patterns, the objective of an intelligent crop recommendation system includes supporting climate change adaptation. By considering climate data, predicting weather patterns, and assessing climate-resilient crop varieties, the system provides recommendations that help farmers adapt to climate risks and maintain agricultural productivity in changing environments.

4.Provide Decision Support: An intelligent crop recommendation system serves as a valuable decision support tool for farmers. By providing accurate and timely recommendations, the system assists farmers in making informed decisions related to crop selection, planting, nutrient management, irrigation enabling them to achieve better outcomes and optimize their farming operations.

5.Facilitate Knowledge Transfer: Another objective is to facilitate knowledge transfer and dissemination of best practices. The system aims to share agricultural knowledge, research findings, and innovative techniques with farmers, empowering them with up-to-date information and helping them adopt modern and efficient farming methods.

iii. Phase wise Proposed Work :

Phase I	Phase II
Finalizing project topic	Coding
Requirement Collection	Database Connectivity
Design the modules	Testing and Validation
Database Design	

2. Facilities Available

(a) Internet.

(b) Required configured machine.

(c) Required software

ABSTRACT

Agriculture and its allied sectors are undoubtedly the largest providers of livelihoods in rural India. Theagriculture sector is also a significant contributor factor to the country's Gross Domestic Product (GDP).Blessing to the country is the overwhelming size of the agricultural sector. However, regrettable is the yield per hectare of crops in comparison to international standards. This is one of the possible causes fora higher suicide rate among marginal farmers in India. This Project proposes a viable and user- friendly yield prediction system for the farmers. The proposed system provides connectivity to farmers. The user provides the area & soil type as input. Machine learning algorithms allow choosing the most profitable crop list or predicting the crop yield for a user-selected crop. To predict the crop yield, selectedMachine Learning algorithms such as Decision Tree, Support Vector Machine (SVM), Gaussian Naïve Bayes, Random Forest (RF), Logistic Regression, and Xgboost are used. Among them, the Random Forest showed the best results with 97% accuracy.

Agriculture is a major contributor to the Indian economy. The common problem existing among the Indian farmers are they don't choose the right cropbased on their soil requirements. Due to this they face a serious setback in productivity. This problem ofthe farmers has been addressed through precision agriculture. Precision agriculture isa modern farming technique that uses research data of soil characteristics, soil types, crop yield data collection and suggeststhe farmers the right crop based on their site- specific parameters. This reduces the wrong choice on a crop and increases the productivity. In this project, we are building an intelligent system, which intendsto assist the Indian farmers in making an informed decision about which crop to grow depending on the sowing season, his farm's geographical location and soil characteristics. Further the system will also provide the farmer, the yield prediction if he plants the recommended crop.

CONTENTS

Abbreviations

IoT: Internet Of Things
CSM: Crop Selection Method
AUC-ROC: Area under the receiver operating characteristic curve
API: Application Programming Interface
HTML: Hyper Text Markup Language
CSS: Cascade Style Sheet
SRS: Software Requirements Specification
ER: Entity-Relationship
DFD: Data Flow Diagram
SQL: Structured Query Language
AI: Artificial Intelligence
nDArray: n-Dimentional Array
PIL: Python Imaging Library
HTTP: Hypertext Transfer Protocol
OS: Operating System
VL: Very Low
VH: Very High
SVM: Support Vector Machine

List Of Figures

List of Tables

Chapter 1

INTRODUCTION

1.1 Introduction:

Agriculture is a significant area for the Indian economy and human survival. It is one of the primary occupations which are essential for human life. It likewise contributes a huge part to our day-to-day life [1]. In most cases farmer's commit suicide due to production loss because they are not able to pay the bank loans taken for farming purposes[12]. We have noticed in present times that the climate is changing persistently which is harmful to the crops and leading farmers towards debt and suicide [18]. These risks can be minimized when various mathematical or statistical methods are applied to data andby using these methods, we can recommend the best crop to the farmer for his Agricultural land so thatit helps him to get maximum profit [12]." Nowadays agriculture has developed a lot in India. "site- specific" farming is the key to Precision agriculture. Although precision agriculture has achieved better enhancements it is still facing certain issues. Precision agriculture plays an important role in the recommendation of crops. The recommendation of crops is dependent on various parameters." Precisionagriculture focuses on identifying these parameters in a site-specific way to identify issues. Not all the results given by precision agriculture are accurate to result but in agriculture, it is significant to have accurate and precise recommendations because, in case of errors, it may lead to heavy material and capital loss. Many research works are being carried out, to attain an accurate and more efficient modelfor crop prediction [11]."

Machine Learning focuses on the algorithm like supervised, unsupervised, and Reinforcement learningand each of them has its advantages and disadvantages. Supervised learning the algorithm assembles amathematical model from a set of data that contains both the inputs and the desired outputs. An unsupervised learning-the algorithm constructs a mathematical model from a set of data that contains only inputs and no desired output labels. Semi-supervised learning- algorithms expand mathematical models from incomplete. This project aims to recommend the most suitable crop based on input parameters like Nitrogen (N), Phosphorous

(P), Potassium (K), PH value of soil, Humidity, Temperature, and Rainfall. This paper predicts the accuracy of the future production of eleven different crops such as rice, maize, chickpea, kidney beans, pigeon peas, moth beans, Mungbam, black gram, lentils, pomegranate, banana, mango, grapes, watermelon, muskmelon, apple, orange, papaya, coconut, cotton, jute, and coffee crops using various supervised machine learning approaches in India and recommends the most suitable crop.

The dataset contains various parameters like Nitrogen (N), Phosphorous (P), Potassium (K), PH value of soil, Humidity,Temperature, and Rainfall. This proposed system applied different kinds of Machine Learning algorithms like Decision Trees, Naïve Bayes (NB), Support Vector Machine (SVM), Logistic Regression, Random Forest (RF), and XGBoost [12]

1.2 Need Of Work :

In today's world, an intelligent crop recommendation system using machine learning is increasingly necessary due to the following reasons:

1. Growing Global Population: The world's population is continuously increasing, leading to higher food demands. To meet the rising food requirements, agricultural systems need to be more efficient and productive. Intelligent crop recommendation systems can assist in maximizing crop yields, ensuring food security for the growing population.
2. Climate Change: Climate change has a significant impact on agriculture, affecting weather patterns, precipitation levels, and the spread of pests and diseases. An intelligent crop recommendation system can analyze historical and real-time climate data to suggest crops that are more resilient to changing environmental conditions, helping farmers adapt to climate change challenges.
3. Data Availability and Accessibility: With the advancement of technology, vast amounts of agricultural data are now available, including weather data, soil information, and crop performance records. An intelligent crop recommendation system can leverage this data to provide accurate and tailored recommendations to

farmers, utilizing the power of machine learning algorithms to extract meaningful insights.

4. Sustainable and Precision Agriculture: The global focus on sustainable practices and precision agriculture has intensified. Intelligent crop recommendation systems align perfectly with these objectives by promoting optimal resource utilization, reducing environmental impact, and enhancing the overall efficiency of agricultural practices.
5. Farmer Empowerment: An intelligent crop recommendation system empowers farmers with data-driven insights, enabling them to make well-informed decisions. This empowerment helps farmers optimize their yields, reduce risks, and improve their economic outcomes.
6. Adoption of Smart Farming Technologies: The agriculture sector is embracing smart farming technologies, including IoT devices, drones, and remote sensing. These technologies generate vast amounts of data, making it necessary to employ ML techniques to process and interpret the information effectively.
7. Addressing Global Challenges: With ongoing challenges like food security, climate change adaptation, and sustainable development, intelligent crop recommendation systems can contribute to global efforts aimed at achieving these objectives.

1.3 Objectives:

- To build a robust model to give correct and accurate prediction of crop sustainability in each statefor the particular soil type and climatic conditions.
- Provide recommendation of the best suitable crops in the area so that the farmer does notincur anylosses.
- Provide profit analysis of various crops based on previous year’s data.
- To recommend optimum crops to be cultivated by farmers based on several parameters and help themmake an informed decision before cultivation

Chapter 2

LITERATURE SURVEY

- Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique by Kumar et al. This paper proposed a method named Crop Selection Method (CSM) to solve crop selection problem, and maximize net yield rate of crop over season and subsequently achieves maximum economic growth of the country. The proposed method may improve net yield rate of crops.

- Agro Consultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms by Zeel et al, This paper proposed and implemented an intelligent crop recommendation sys- tem, which can be easily used by farmers all over India. This system would assist the farmers in making an informed decision about which crop to grow depending on a variety of environmental and geographical factors. We have also implemented a secondary system, called Rainfall Predictor, which predicts the rainfall of the next 12 months.

- Development of Yield Prediction System Based on Real-time Agricultural meteorological Information Haedong et al. This paper contains about the research and the building of an effective agricultural yield forecasting system based on real-time monthly weather. It is difficult to predict the agricultural crop production because of the abnormal weather that happens every year and rapid regional climate change due to global warming. The development of agricultural yield forecasting system that leverages real-time weather information is urgently required. In this research, we cover how to process the number of weather Intelligent data (monthly, daily) and how to configure the prediction system. We establish a non-parametric statistical model on the basis of 33 years of agricultural weather in- formation. According to the implemented model, we predict final production using the monthly weather information. This paper contains the results of the simulation.

- Analysis of Soil Behaviour and Prediction of Crop Yield using Data Mining Approach Monali et alComputer science and Engineering GGITS, Jabalpur. This work presents a system, which uses datamining techniques in order to predict the category of the analyzed soil datasets. The category, thus predicted will indicate the yielding of crops. The problem of predicting the crop yield is formalizedas a classification rule, where Naive Bayes and KNearest Neighbor methods are used.

- Crop Recommendation System for Precision Agriculture S.Pudumalar* et al, This paper, proposes arecommendation system through an ensemble model with majority voting technique using Randomtree, CHAID, K-Nearest Neighbor and Naive Bayes as learners to recommend a crop for the site specific parameters with high accuracy and efficiency.

2.1 Existing Systems :

Several existing intelligent crop recommendation systems already utilize machine learning and data analysis techniques. Here are a few notable examples:

1. IBM Watson Decision Platform for Agriculture: IBM's Watson Decision Platform for Agriculture combines IoT sensors, weather data, and ML algorithms to deliver personalized crop recommendations. The platform integrates data from soil moisture sensors, drones, and other sources to provide insights on irrigation scheduling, disease prediction, and optimal planting times.

2. FarmBeats: FarmBeats is a Microsoft project that aims to empower smallholder farmers through ML-based crop recommendations. It utilizes IoT sensors, aerial imagery, and ML algorithms to generate actionable insights, such as crop water requirements, fertilizer recommendations, and yield predictions.

3. AgroPilot: AgroPilot is an intelligent agriculture system that employs ML algorithms to analyze data from soil sensors, weather stations, and satellite imagery. It provides

farmers with real-time recommendations on irrigation, fertilization, and crop protection to optimize crop yields and reduce resource waste.

4. Taranis: Taranis is an AI-powered precision agriculture platform that uses ML algorithms to analyze aerial imagery and satellite data. It provides farmers with recommendations for pest and disease management, crop health monitoring, and yield prediction, helping optimize crop production and reduce chemical inputs.

2.2 Problem Definition :

The problem that an intelligent crop recommendation system aims to address is the challenge faced by farmers in selecting the most suitable crops for their land using traditional and non-scientific methods. This problem is particularly significant in countries where a significant portion of the population is engaged in farming, and access to accurate and up-to-date information is limited.

The existing methods of crop selection often rely on outdated practices, personal experiences, or anecdotal information, which can lead to suboptimal crop choices. This can result in lower yields, increased vulnerability to weather fluctuations and pests, and reduced profitability for farmers.

Additionally, the lack of access to scientific research and case studies specific to developing countries further hinders the ability of farmers to make informed decisions. This creates a gap in knowledge and prevents the adoption of best practices and optimized crop selection techniques.

Therefore, the problem can be summarized as follows:

- Farmers face challenges in determining the best suited crops for their land using traditional and non-scientific methods.
- The absence of research and case studies for developing countries further exacerbates the problem.

- Sub optimal crop choices can lead to lower yields, increased vulnerability, and reduced profitability for farmers.
- Limited access to accurate and up-to-date information impedes effective decision-making.

2.3 Proposed System:

2.3.1 System Architecture:

A system architecture is a conceptual model using which we can define the structure and behavior of that system. It is a formal representation of a system. Depending on the context, system architecture can be used to refer to either a model to describe the system or a method used to build the system. Building a proper system architecture helps in analysis of the project, especially in the early stages. Depicts the system architecture and is explained in the following section.

- Machine Learning: Machine learning is a subfield of artificial intelligence that focuses on developing algorithms and models that enable computers to learn from data and make predictions or decisions without explicit programming. It involves techniques such as supervised learning, unsupervised learning, and reinforcement learning.
- Supervised Learning: Supervised learning is a machine learning technique where a model learns from labeled training data. The model is trained on input-output pairs, where the inputs represent the features (e.g., weather, soil conditions), and the outputs represent the desired prediction (e.g., recommended crop). Popular supervised learning algorithms include decision trees, random forests, support vector machines, and neural networks.
- Decision Trees: Decision trees are predictive models that map observations about an item to conclusions about the item's target value. They are built by recursively partitioning the data based on feature values, resulting in a tree-like structure where each internal node represents a decision based on a feature, and each leaf node represents a class or prediction.

• Reinforcement Learning: Reinforcement learning is a branch of machine learning where an agent learns to make decisions and take actions in an environment to maximize a cumulative reward. In the context of crop recommendation, reinforcement learning techniques can be used to optimize resource allocation, irrigation scheduling, or pest control strategies to maximize crop yield.

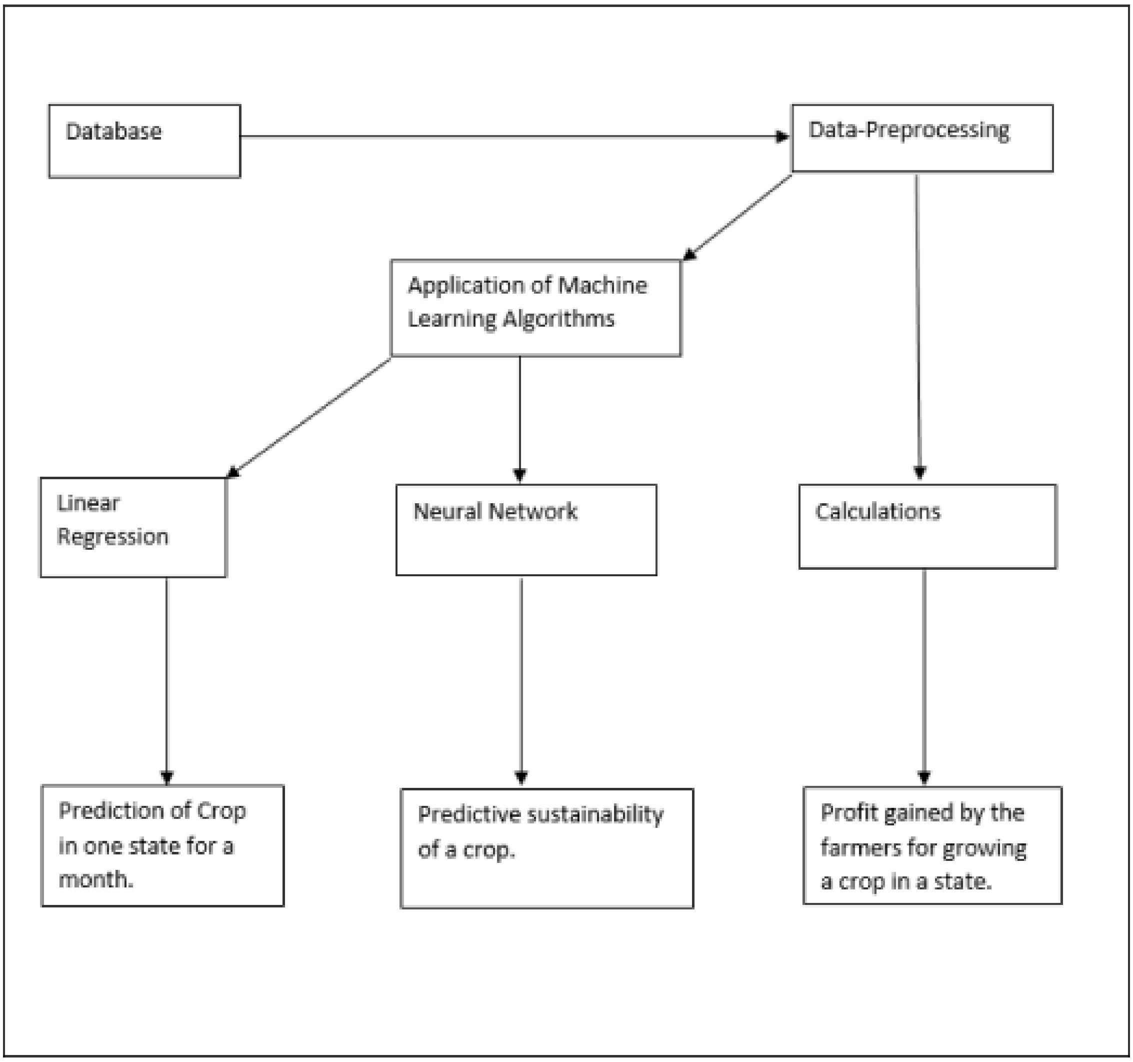

Figure 2.1: system Architecture

- **Data Collection**

  Our information includes the temperature, N, P, K, humidity, ph, and rainfall as an attribute, as wellas the outcomes of crops that may be cultivated in that soil type. The dataset consists of a few majorcrops which are mostly cultivated as wheat, sugarcane, rice, maize, chickpea, kidney-beans, pigeon- peas, moth-beans, mung-bean, blackgram, lentil, pomegranate, banana, mango, grapes, watermelon, muskmelon, apple, orange, papaya, coconut, cotton, jute, coffee.

- **Data Preprocessing**

  Data preprocessing is a data mining approach that entails converting raw data into a format that can be understood. Because the original dataset may have a large number of missing values, all of them should be eliminated at first. Missing values are represented by a dot in the dataset, and their existence can degrade the overall value of the data as well as impair performance. As a result, we replace these numbers with the mean values to fix this problem. The second step is to create the classlabels. Because we're utilizing supervised learning, there should be a class label for each entry in the dataset, which is produced during the preprocessing phase

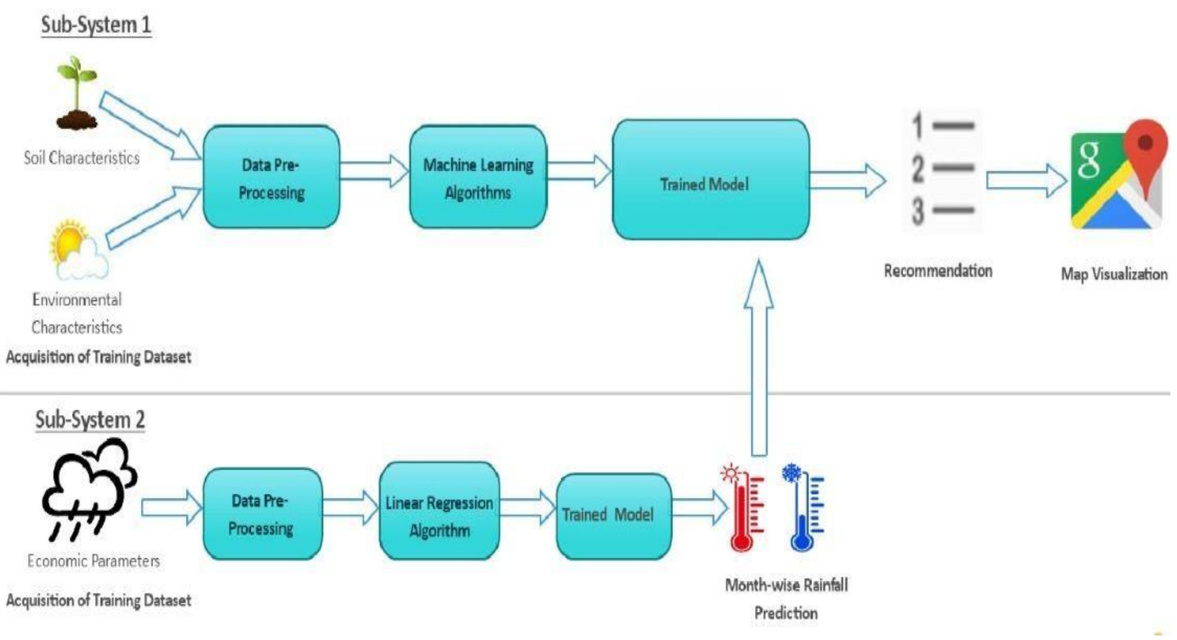

Figure 2.2: Data Preprocess

- **Regression Analysis**

  Regression Analysis is a predictive modelling technique that examines the relationship between a dependent or target variable and an independent or predictor variable. It covers linear, multiple linear, and non-linear regression models, among others. Simple linear regression is the most used model. Polynomial regression is a type of regression method in which the link between the independent variable x and the dependent variable y.

  polynomial regression fits a nonlinear relationship between the value of x and the associated conditional mean of y, denoted by E(y |x). Despite the fact that polynomial regression fits a nonlinearmodel to the data, as a statistical estimation problem it is linear in the sense that in the unknown parameters inferred from the data, the regression function E(y |x) is linear. As a result, polynomial regression is regarded as a subset of multiple linear regressions. The predicted value of y may be modelled as an nth degree.

  - Ensemble Learning: Ensemble learning combines multiple machine learning models to make more accurate predictions. It aggregates the outputs of individual models, such as decision trees, and combines them to obtain a final prediction. Examples of ensemble learning techniques include bagging (e.g., random forests) and boosting (e.g., gradient boosting).
  - Feature Selection: Feature selection is the process of selecting a subset of relevant features from a larger set of available features. It aims to reduce dimensionality, improve model performance, and eliminate irrelevant or redundant features. Techniques for feature selection include filter methods, wrapper methods, and embedded methods.
  - Model Evaluation: Model evaluation is performed to assess the performance and generalization capability of the trained models. Common evaluation metrics include accuracy, precision, recall, F1-score, area under the receiver operating characteristic curve (AUC-ROC), and mean squared error (MSE). Cross-validation

and holdout validation are commonly used techniques for evaluating model performance.

- Real-time Data Integration: Real-time data integration involves the continuous collection and integration of real-time data from various sources such as weather APIs, satellite imagery, and IoT sensors. This data is incorporated into the crop recommendation system to provide up-to-date information for improved predictions and recommendations.
- Model Explainability: Model explainability refers to the ability to understand and interpret the decisions and predictions made by machine learning models. Interpretability is important in agriculture, as farmers need to trust and comprehend the recommendations provided by the system. Techniques such as feature importance analysis, decision rule extraction, or model-agnostic interpretability methods can be employed to enhance the transparency and explainability of the crop recommendation models.
- Continuous Learning: Continuous learning refers to the ability of a system to adapt and update its knowledge continuously based on new data. In the context of crop recommendation, continuous learning can be utilized to incorporate new research findings, updated crop models, or real-time data into the system. This ensures that the recommendations remain up to date and accurate over time.

2.3.2 Modules to be Implement:

Front-End Development:

- Design and develop the user interface of the webpage using HTML, CSS, and JavaScript. Create a visually appealing and user-friendly layout that allows farmers to interact with the system.
- Implement input forms for farmers to enter their land and environmental data, such as soil conditions and location.
- Incorporate interactive elements like dropdown menus, sliders, or checkboxes to capture user preferences.

- Display the generated crop recommendations in a clear and organized manner, possibly using tables, charts, or visualizations.
- Ensure the webpage is responsive and compatible with different devices and screen sizes.

Back-End Development:

- Set up a server-side scripting language (e.g., Python, Node.js) to handle data processing and communication between the front-end and back-end components.
- Develop server-side logic to receive user inputs from the webpage and pass them to the recommendation engine for processing.
- Implement data validation and error handling to ensure the integrity of user inputs and provide meaningful error messages if necessary.
- Integrate the recommendation engine module with the back-end to generate crop recommendations based on the received user data.
- Retrieve the recommended crop information from the recommendation engine and send it back to the front-end for display.

2.4 Advantageous of Proposed System :

- Improved Crop Selection: The system leverages machine learning algorithms to analyze a wide range of data, including weather patterns, soil conditions, historical records, and crop characteristics. By considering these factors, the system can provide more accurate and informed recommendations for farmers, leading to improved crop selection. This can result in higher crop yields, reduced losses, and increased profitability for farmers.
- Increased Efficiency: The system automates the crop recommendation process, reducing the time and effort required for farmers to manually research and analyze data. By providing real-time and personalized recommendations, farmers can make decisions more efficiently, enabling them to focus on other aspects of their farming operations.
- Enhanced Decision-Making: The system provides predictive insights and

recommendations based on machine learning models trained on essential environmental and economic parameters. By leveraging these insights, farmers can make data-driven decisions about crop selection, resource allocation, and risk management. This leads to more informed and strategic decision-making, minimizing the potential for errors or suboptimal choices.

- Flexibility and Adaptability: The system can be customized to accommodate different farming conditions, including variations in soil types, climates, and geographical regions. By considering diverse land conditions and accurately identifying suitable crops, the system ensures its recommendations are tailored to specific farming contexts. It can adapt to changing conditions, incorporate new data sources, and continuously improve its recommendations over time.
- Increased Productivity and Profitability: By recommending crops with higher potential yields and economic viability, the system can help farmers maximize their productivity and profitability. Farmers can optimize resource allocation, reduce the risk of crop failure, and explore new crop options based on the system's recommendations. This can lead to improved financial outcomes and a more sustainable agricultural sector.
- Accessible and User-Friendly: The system can be accessed through a user-friendly web interface, making it easily accessible to farmers with minimal technical expertise. The interface provides a straightforward means of inputting data, receiving recommendations, and visualizing relevant information. This accessibility enables farmers from diverse backgrounds to benefit from the system's capabilities.
- Sustainability and Environmental Impact: The system considers environmental parameters in its recommendations, such as water availability, soil conservation, and climate suitability.

Chapter 3

SYSTEM ANALYSIS AND DESIGN

3.1 Requirements specification:

A software requirements specification (SRS) is a description of a software system to be developed. It lays out functional and non-functional requirements, and may include a set of use cases that describe user interactions that the software must provide. In order to fully understand one"s project, it is very important that they come up with a SRS listing out their requirements, how are they going to meet it and how will they complete the project. It helps the team to save upon their time as they are able to comprehend how are going to go about the project. Doing this also enables the team to find out about the limitations and risks early on. Requirement is a condition or capability to which the system must conform. Requirement Management is a systematic approach towards eliciting, organizing and documenting the requirements of the system clearly along with the applicable at- tributes. The elusive difficulties of requirements are not always obvious and can come from any number of sources.

3.1.1 Non-Functional Requirements:

Nonfunctional requirements are the requirements which are not directly concerned with the specific function delivered by the system. They specify the criteria that can be used to judge the operation of a system rather than specific behaviors. They may relate to emergent system properties such as reliability, response time and store occupancy. Nonfunctional requirements arise through the user needs, because of budget constraints, organizational policies and the need for interoperability with other software and hardware systems.

➢ **Organizational Requirements Process Standards:**

The standards are used to develop the application which is the standard used by the developers. Design Methods: Design is one of the important stages in the software engineering process. This stage is the first step in moving from problem to the solution domain. In other words, starting with what is needed design takes us to work how to satisfy the needs.

➢ **Basic Operational Requirements**:

The customers are those that perform the eight primary functions of systems engineering, with special emphasis on the operator as the key customer. Operational requirements will define the basic need and, at a minimum, will be related to these following points:

- Mission profile or scenario: It describes about the procedures used to accomplish mission objective. It also finds out the effectiveness or efficiency of the system.
- Performance and related parameters: It points out the critical system parameters to accomplish the mission.
- Utilization environments: It gives a brief outline of system usage. Finds out appropriate environments for effective system operation.
- Operational life cycle: It defines the system lifetime.

➢ **System Configuration Hardware System Configuration:**

- Processor: 2 gigahertz (GHz) or faster processor or SoC
- RAM: 6 gigabyte (GB) for 32-bit or 8 GB for 64-bit
- Hard disk space: =16GB.

➢ **Software Configuration:**

- Operating System: Windows XP/7/8/8.1/10, Linux and Mac
- Coding Language: Python
- Tools:

    1. Python
    2. HTML
    3. CSS
    4. JavaScript
    5. Pandas
    6. Numpy
    7. Tensorflow

8. Keras

9. Sickitlearn

3.1.2 Functional Requirements:

Functional Requirement defines a function of a software system and how the system must behave when presented with specific inputs or conditions. These may include calculations, data manipulation and processing and other specific functionality. Following are the functional requirements on the system:

1. All the data must be in the same format as a structured data.
2. The data collected will be vectored and sent across to the classifier.

3.2 UML Diagrams:

3.2.1 ER Diagram:

An Entity-Relationship (ER) diagram is a visual representation of the entities (objects or concepts) within a system and the relationships between them. It is a widely used modeling tool in software engineering and database design.

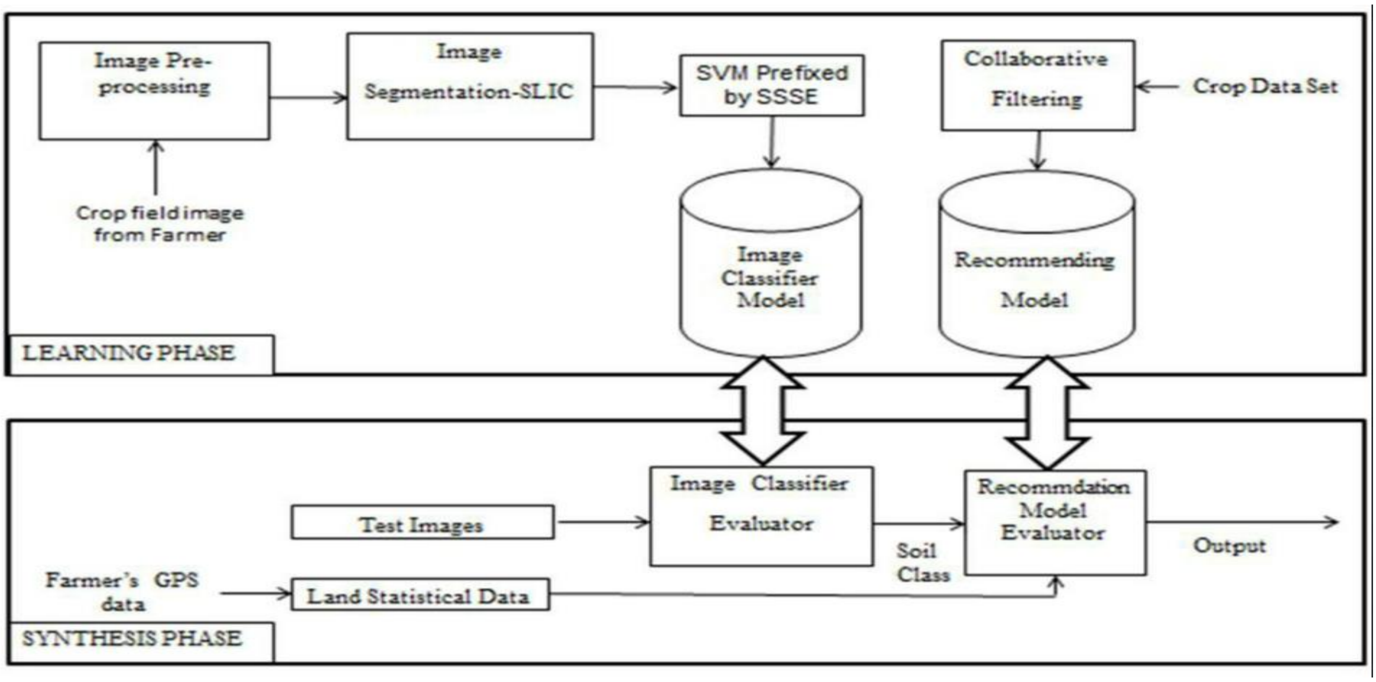

Figure 3.1: ER Diagram

3.2.2 Data Flow Diagram :

A Data Flow Diagram (DFD) is a structured analysis and design tool that can be used for flow charting. A DFD is a network that describes the flow of data and the processes that change or transform the data throughout a system. This network is constructed by using a set of symbols that do not imply any physical implementation. It has the purpose of clarifying system requirements and identifying major transformations. So, it is the starting point of the design phase that functionally decomposes the requirements specifications down to the lowest level of detail. DFD can be considered to an abstraction of the logic of an information-oriented or a process-oriented system flow-chart. For these reasons DFDs are often referred to as logical data flow diagrams.

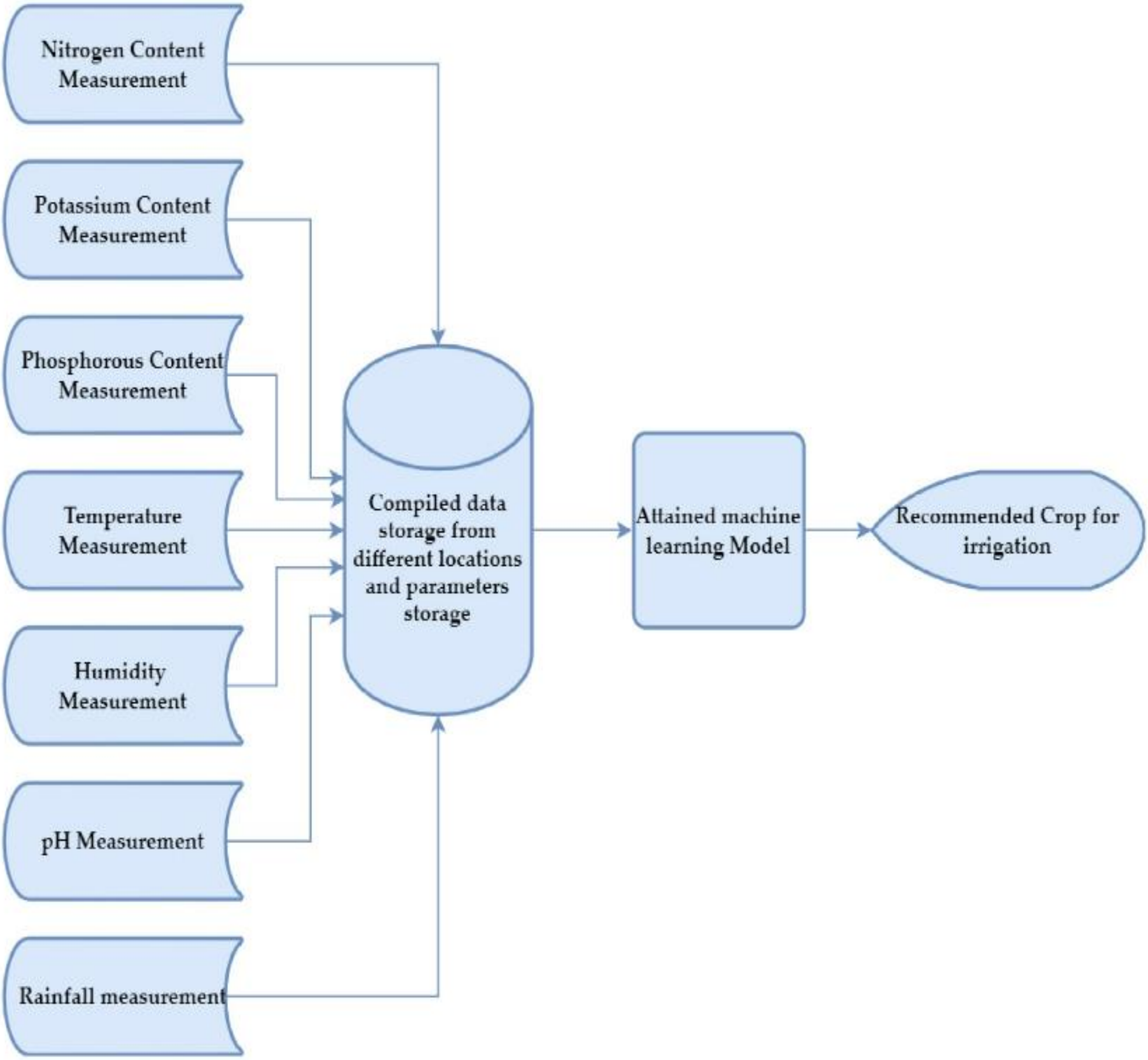

Figure 3.2: Data Flow Diagram

3.3 Design and Test Steps / Criteria:

Design

3.3.1 System Development Methodology:

System Development methodology is the development of a system or method for a unique situation. Having a proper methodology helps us in bridging the gap between the problem statement and turning it into a feasible solution. It is usually marked by converting the System Requirements Specifications (SRS) into a real world solution. System design takes the following inputs:

- Statement of work.
- Requirement determination plan.
- Current situation analysis.
- Proposed system requirements including a conceptual data model and metadata (data about data).

3.3.2 Model Phases:

The waterfall model is a sequential software development process, in which progress is seen as owing steadily downwards (like a waterfall) through the phases of Requirement initiation, Analysis, Design, Implementation, Testing and maintenance.

Requirement Analysis: This phase is concerned about collection of requirement of the system. This process involves generating document and requirement review.

System Design: Keeping the requirements in mind the system specifications are translated in to a software representation. In this phase the designer emphasizes on:- algorithm, data structure, software architecture etc.

Coding: In this phase programmer starts his coding in order to give a full sketch of product. In other words system specifications are only converted in to machine.

Implementation: The implementation phase involves the actual coding or programming of the software. The output of this phase is typically the library, executable, user manuals and additional soft-ware documentation.

Testing: In this phase all programs (models) are integrated and tested to ensure that the complete system meets the software requirements. The testing is concerned with verification and validation.

Maintenance: The maintenance phase is the longest phase in which the software is updated to fulfill the changing customer needs, adapt to accommodate changes in the external environment, correct errors and oversights previously undetected in the testing phase, enhance the efficiency of the software.

3.3.3 Sequence Diagram :

A Sequence diagram is an interaction diagram that shows how processes operate with one another and what is their order. It is a construct of a Message Sequence Chart. A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called event diagrams or event scenarios. Sequence diagram are an easy and intuitive way of describing the behavior of a system by viewing the interaction between the system and the environment. A sequence diagram shows an interaction arranged in a time sequence. A sequence diagram has two dimensions: vertical dimension represents time; the horizontal dimension represents the objects existence during the interaction.

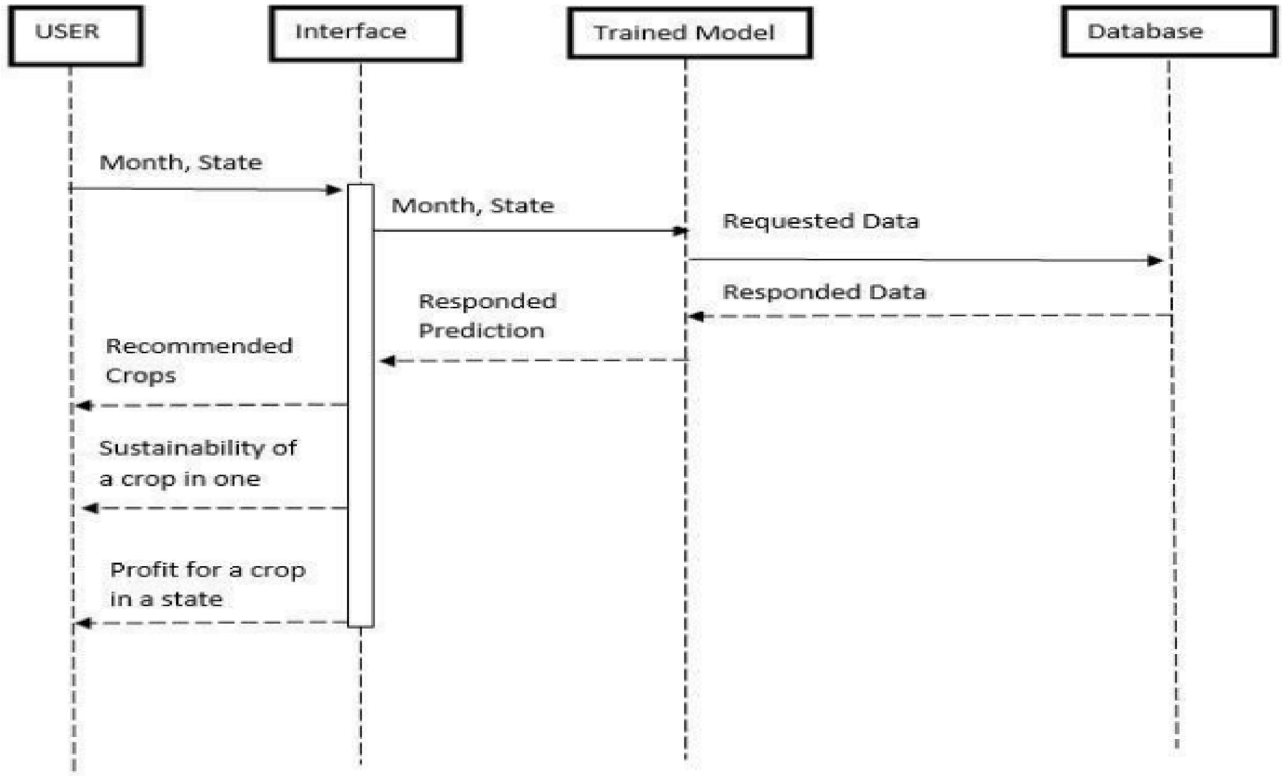

Figure 3.3: Sequence Diagram

Test Steps

Here are some examples of test cases for an intelligent crop recommendation system:

1. **Test Case: Input Validation**
    - Verify that the system handles invalid or unexpected input gracefully.
    - Test with empty or null inputs for required fields.
    - Test with input values outside the expected ranges or formats.
    - Ensure appropriate error messages or notifications are displayed.
2. **Test Case: Crop Recommendation Accuracy**
    - Provide input data for a specific farming scenario.
    - Verify that the system recommends suitable crops based on the given data.
    - Cross-reference the recommendations with expert knowledge or historical data to validate accuracy.

3. **Test Case: Boundary Conditions**
    - Test with extreme values for environmental factors, such as very high or low temperatures, rainfall, or soil pH.
    - Verify that the system handles these boundary conditions correctly and provides appropriate recommendations.
4. **Test Case: Performance and Scalability**
    - Test the system's response time with a large dataset and multiple concurrent users.
    - Measure the time taken for the system to process and deliver recommendations.
    - Monitor resource utilization, such as CPU and memory usage, during peak loads.
5. **Test Case: User Interface and Usability**
    - Verify that the user interface is intuitive, visually appealing, and easy to navigate.
    - Test various user interactions, such as entering data, viewing recommendations, and changing preferences.
    - Assess the system's responsiveness and ensure a smooth user experience.
6. **Test Case: Integration Testing**
    - Test the integration of the system with external data sources, such as weather APIs or market data providers.
    - Verify that the system retrieves and processes data from these sources accurately and in a timely manner.
7. **Test Case: Security Testing**
    - Validate the system's security measures, such as user authentication and access controls.
    - Test for potential vulnerabilities, such as SQL injection or cross-site scripting attacks.
    - Ensure that sensitive user data is handled securely and encrypted during transmission.

8. **Test Case: Error Handling**
   - Test scenarios where the system encounters errors or exceptions.
   - Verify that error messages are displayed appropriately and provide meaningful information for troubleshooting.
9. **Test Case: Regression Testing**
   - Re-run previously executed test cases after making changes or bug fixes.
   - Verify that the system's existing functionality has not been impacted by recent modifications.
10. **Test Case: Compatibility Testing**
    - Test the system's compatibility with different web browsers, operating systems, or devices.
    - Verify that the user interface and functionality work as expected across different environments.

3.4 Testing Process:

Data Analysis :

One of the first steps we perform during implementation is an analysis of the data. This was done by us in an attempt to find the presence of any relationships between the various attributes present in the dataset.

Acquisition of Training Dataset: The accuracy of any machine learning algorithm depends on the number of parameters and the correctness of the training dataset. We In this project analyzed multiple datasets collected from Government website -https://data.gov.in/ and Kaggle and carefully selected the parameters that would give the best results. Many work done in this field have considered environmental parameters to predict crop sustainability some have used yield as major factor where as in some works only economic factors are taken into consideration. We have tried to combine both environmental parameters like rainfall, temperature, ph, nutrients in soil, soil type and economic parameters like production, and yield to provide accurate and reliable recommendation to the farmer on which crop is suitable.

	A	B	C	D	E	F	G	H
1	N	P	K	temperatu	humidity	ph	rainfall	label
2	90	42	43	20.8797	82.0027	6.50299	202.936	rice
3	85	58	41	21.7705	80.3196	7.0381	226.656	rice
4	60	55	44	23.0045	82.3208	7.84021	263.964	rice
5	74	35	40	26.4911	80.1584	6.9804	242.864	rice
6	78	42	42	20.1302	81.6049	7.62847	262.717	rice
7	69	37	42	23.058	83.3701	7.07345	251.055	rice
8	69	55	38	22.7088	82.6394	5.70081	271.325	rice
9	94	53	40	20.2777	82.8941	5.71863	241.974	rice
10	89	54	38	24.5159	83.5352	6.68535	230.446	rice
11	68	58	38	23.224	83.0332	6.33625	221.209	rice
12	91	53	40	26.5272	81.4175	5.38617	264.615	rice
13	90	46	42	23.979	81.4506	7.50283	250.083	rice
14	78	58	44	26.8008	80.8868	5.10868	284.436	rice
15	93	56	36	24.015	82.0569	6.98435	185.277	rice
16	94	50	37	25.6659	80.6639	6.94802	209.587	rice
17	60	48	39	24.2821	80.3003	7.0423	231.086	rice
18	85	38	41	21.5871	82.7884	6.24905	276.655	rice
19	91	35	39	23.7939	80.4182	6.97086	206.261	rice
20	77	38	36	21.8653	80.1923	5.95393	224.555	rice
21	88	35	40	23.5794	83.5876	5.85393	291.299	rice
22	89	45	36	21.325	80.4748	6.44248	185.497	rice
23	76	40	43	25.1575	83.1171	5.07018	231.384	rice
24	67	59	41	21.9477	80.9738	6.01263	213.356	rice
25	83	41	43	21.0525	82.6784	6.25403	233.108	rice
26	98	47	37	23.4838	81.3327	7.37548	224.058	rice

Table 3.1: complete dataset

Data Preprocessing:

After analyzing and visualizing the data, the next step is preprocessing. Data preprocessing is an important step as it helps in cleaning the data and making it suitable for use in machine learning algorithms. Most of the focus in preprocessing is to remove any outliers or erroneous data, as well as handling any missing values. Missing data can be dealt with in two ways. The first method is to simply remove the entire row which contains the missing or erroneous value. While easy to execute method, it is better to use online large datasets. Using this method on small datasets can reduce the dataset size too much, especially if there are a lot of missing values. This can severely affect the accuracy of the result. Since, ours is a relatively small dataset, we will not be using this method. The dataset that we used had values that were in string format so we had to transform and encode into the integer valued so as to pass as an input to the neural network. First, we converted the data into pandas categorical data and then generated codes for crops and states respectively we than appended these and created separated datasets. Further to reduce the amount of data going into the linear regression model we filtered the crops based on the required nutrients and nutrients present in the soil. If the nutrient content of the soil was below that required by the crops, then that crop was discarded, in this way we were able to reduce the training time a lot.

Chapter 4

METHODOLOGY

4.1 Plan of Implementation:

The steps involved in this system implementation are:-

a) **Acquisition of Training Dataset:** The accuracy of any machine learning algorithm depends on the number of parameters and the correctness of the training dataset. For the system, we are using various datasets all downloaded for government website and kaggle.

Datasets include:-

Cost of cultivation per dataset for major crops in each state

Yield dataset

Modal price of crops Standard price of crops

Soil nutrient content dataset Rainfall Temperature dataset

b) Data Preprocessing: This step includes replacing the null and 0 values for yield by - 1 so that it does not affect the overall prediction. Further we had to encode the dataset so that it could be fed into the neural network.

c) Training model and crop recommendation: After the preprocessing step we used the data-set to train different machine learning models like neural network and linear regression to attain accuracy as high as possible.

Theoretical Background

4.2 Overview on Machine Learning:

Machine learning is an application of artificial intelligence (AI) that gives systems the ability to automatically learn and evolve from experience without being specially programmed by the programmer. The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The main aim of machine learning is to allow computers to learn automatically and adjust their actions to improve the accuracy and usefulness of the program, without any human intervention or assistance. Traditional writing of programs for a computer can be defined as automating the procedures to be performed on input data in order to create output artifacts. Almost always, they are linear, procedural and logical. A traditional program is written in a programming language to some specification, and it has properties like:

• We know or can control the inputs to the program.

• We can specify how the program will achieve its goal.

• We can map out what decisions the program will make and under what conditions it makes them.

• Since we know the inputs as well as the expected outputs, we can be con dent that the program will achieve its goal

Traditional programming works on the premise that, as long as we can define what a program needs to do, we are confident we can define how a program can achieve that goal. This is not always the case as sometimes; however, there are problems that you can represent in a computer that you cannot write a traditional program to solve. Such problems resist a procedural and logical solution. They have properties such as:

• The scope of all possible inputs is not known beforehand.

• You cannot specify how to achieve the goal of the program, only what that goalis.

• You cannot map out all the decisions the program will need to make to achieveits goal.

• You can collect only sample input data but not all possible input data for the program.

4.2.1. Supervised and Unsupervised Learning

Machine learning techniques can be broadly categorized into the following types:

Supervised learning takes a set of feature/label pairs, called the training set. From this training set the system creates a generalized model of the relationship between the set of descriptive features and the target features in the form of a program that contains a set of rules. The objective is to use the output program produced to predict the label for a previously unseen, unlabeled input set of features, i.e. to predict the outcome for some new data. Data with known labels, which have not been included in the training set, are classified by the generated model and the results are compared to the known labels. This dataset is called the test set. The accuracy of the predictive model can then be calculated as the proportion of the correct predictions the model labeled out of the total number of instances in the test set.

Unsupervised learning takes a dataset of descriptive features without labels as a train- ing set. In unsupervised learning, the algorithms are left to themselves to discover interesting structures in the data. The goal now is to create a model that finds some hidden structure in the dataset, such as natural clusters or associations. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system does not figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data. Unsupervised learning can be used for clustering, which is used to discover any inherent grouping that is already present in the data. It can also be used for association problems, by creating rules based on the data and finding relationships or associations between them.

Semi-supervised machine learning falls somewhere in between supervised and unsupervised learning, since they use both labeled and unlabeled data for training typically a small amount of labeled data and a large amount of unlabeled data. The systems that use this method are able to considerably improve learning accuracy. Usually, semi- supervised learning is chosen when the acquired labeled data requires skilled and relevant resources in order to train it / learn from it. Otherwise, acquiring labeled data generally does not require additional resources.

Reinforcement machine learning algorithms is a learning method that interacts with its environment by producing actions and discovers errors or rewards. Machine learning algorithms are tools to automatically make decisions from data in order to achieve some over-arching goal or requirement. The promise of machine learning is that it can solve complex problems automatically, faster and more accurately than a manually specified solution, and at a larger scale. Over the past few decades, many machine learning algorithms have been developed by researchers, and new ones continue to emerge and old ones modified.

4.3 Machine Learning Tools

There are many different software tools available to build machine learning models and to apply these models to new, unseen data. There are also a large number of well-defined machine learning algorithms available. These tools typically contain libraries implementing some of the most popular machine learning algorithms. They can be categorized as follows:

- Pre-built application-based solutions.
- Programming languages which have specialized libraries for machine learning

Using programming languages to develop and implement models is more flexible and gave us better control of the parameters to the algorithms. It also allows us to have a better understanding of the output models produced. Some of the popular programming languages used in the field of machine learning are:

- **Python:** Python is an extremely popular choice in the field of machine learning and AI development. Its short and simple syntax make it extremely easy to learn

- **R:** R is one of the most effective and efficient languages for analyzing and manipulating data in statistics. Using R, we can easily produce well-designed publication-quality plot, including mathematical symbols and formulae where needed. Apart from being a general purpose language, R has numerous of pack- ages like RODBC, Gmodels, Class and Tm which are used in the field of ma- chine learning. These packages make the implementation of machine learning algorithms easy, for cracking the business associated problems

• **Tensorflow:** TensorFlow is an end-to-end open source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries, and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy MLpowered applications. TensorFlow was originally developed by researchers and engineers working on the Google Brain team within Google's Machine Intelligence Research organization to conduct machine learning and deep neural networks research. The system is general enough to be applicable in a wide variety of other domains, as well. TensorFlow provides stable Python and C++ APIs, as well as non-guaranteed backward compatible API for other languages.

4.4 Libararies:

1. Scikit learn:

SciKit learn is an open source machine learning library built for python. Since its release in 2007, Scikit-learn has become one of the most popular open source machine learning libraries. Scikit-learn (also called sklearn) provides algorithms for many ma- chine learning tasks including classification, regression, dimensionality reduction and clustering. The documentation for scikit-learn is comprehensive, popular and well maintained. Sklearn is built on mature Python Libraries such as NumPy, SciPy, and matplotlib. While languages such as R and MATLAB are extremely popular and useful for ma- chine learning, we decided to choose Python along with its SciKit-learn libraries as our programming language of choice. The reasons for this are:

- We already have some familiarity and exposure to Python, and thus have a smaller learning curve.
- Both Python and Scikit-learn have excellent documentation and tutorials avail- able online.
- The number of classic machine learning algorithms that come with Scikit-learn, and the consistent patterns for using the different models i.e., each model can be used with the same basic commands for setting up the data, training themodel and using the model for prediction. This makes it easier to try a range of machine learning algorithms on the same data.
- The machine learning algorithms included with sklearn have modifiable parameters known

as hyper-parameters that effect the performance of the model. These usually have sensible default values, so that we can run them without needing a detailed knowledge or understanding of their semantics.

- The IPython notebook, which is an interactive computational environment for Python, in which a user can combine code execution, rich text, mathematics and plots in a web page. This functionality allows us to provide the notebooks we used to run our experiments almost as an audit and in a presentable.

2. Streamlit:

Streamlit is an open-source Python library used for building interactive web applications and data dashboards. It simplifies the process of creating and sharing data-focused applications by providing a user-friendly and intuitive interface. With Streamlit, you can turn your Python scripts into interactive web apps quickly and easily. Streamlit provides a streamlined and efficient way to build interactive web applications for data analysis, machine learning, and visualization. Its simplicity and ease of use make it a popular choice among data scientists and developers looking to showcase their work or create interactive data-driven experiences.

3. Numpy:

NumPy (Numerical Python) is a powerful Python library for numerical computing. It provides a multidimensional array object, along with a collection of mathematical functions, linear algebra operations, random number generation capabilities, and tools for working with arrays.

Here are some key features and functionalities of NumPy:

N-dimensional array: NumPy's main feature is its ndarray (n-dimensional array) object, which provides efficient storage and manipulation of homogeneous data. Arrays in NumPy can have any number of dimensions and can hold elements of the same data type.

Mathematical functions: NumPy includes a comprehensive collection of mathematical functions for performing operations on arrays. These functions are optimized for efficient numerical computations and can be applied element-wise to arrays or along specified axes. Linear algebra operations: NumPy provides a wide range of linear algebra functions, such as matrix multiplication, eigenvalues and eigenvectors, solving linear equations, and computing matrix inverses. These functions are essential for many scientific and engineering applications.

- Broadcasting: NumPy supports broadcasting, which is a powerful mechanism for performing operations on arrays of different shapes. Broadcasting allows for efficient element-wise operations between arrays without explicitly replicating the data.
- Random number generation: NumPy includes a random module that allows for the generation of random numbers from various probability distributions. This is useful for simulations, statistical analysis, and generating random data for testing and experimentation.

4. **Image(PIL):**

PIL (Python Imaging Library) is a popular library for opening, manipulating, and saving various image file formats in Python. It provides a wide range of functions and methods for performing various image processing tasks.

To use the PIL library in Python, you need to install it first. You can install it using pip, the Python package manager, with the following command:

```
pip install Pillow
```

Once PIL (or its fork, Pillow) is installed, you can import the Image module and use its functions and classes to work with images. Here's an example of how to use PIL to open and display an image:

```python
from PIL import Image

# Open an image file
image = Image.open('example.jpg')

# Display the image
image.show()
```

PIL provides many other functions and methods for different image processing operations. You can refer to the PIL documentation or the Pillow documentation (which is a maintained fork of PIL) for more detailed information on the available functions and their usage.

5. **Pickle:**

   "Pickle" is a Python module that allows you to serialize (convert objects into a byte stream) and deserialize (convert byte stream back into objects) Python objects. It is commonly used for saving and loading objects to/from files or transmitting them over a network.

   The "pickle" module provides two main functions: "pickle. dump()" and "pickle. load()", as well as other related functions.

6. **Request:**

   The requests library is a popular Python module used for making HTTP requests to interact with web servers and retrieve data from web resources. It provides a simple and intuitive interface for sending HTTP requests and handling the responses.

7. **OS:**

   The "OS" library in Python provides functions for interacting with the operating system. It allows you to perform various operating system-related tasks such as file and directory operations, environment variables, process management, and more.

4.5 Dataset

For the system, we are using various datasets all downloaded for government website and kaggle.

Datasets include:-

Soil nutrient content dataset

Rainfall Temperature dataset

A brief description of the datasets:

- Yield Dataset: This dataset contains yield for 16 major crops grown across all the states in kg per hectare Yield of 0 indicates that the crop is not cultivated in the respective state.
- Soil nutrient content dataset: This dataset has five columns with the attributes in the order-State, Nitrogen content, Phosphorous content, Potassium content and average ph. The nutrient content is represented with encoded alphabets- VL,L,M,H,VH with the meaning: VL -Very Low L-Low M-Medium H-High VH-Very high
- Rainfall Temperature dataset: This dataset contains crops, max and min rainfall, max and min temperature, max and min rainfall and ph values.

4.6 Machine Learning Algorithms

Machine Learning algorithms used in the recommendation system are:

- Linear Regression: Linear regression is a linear approach to modeling the relationship between a scalar response (and dependent variable) and one or more explanatory variables (or independent variables). Linear regression is used for finding linear relationship between target and one or more predictors. It fits a linear model with coefficients w = (w1, . . ., wp) to minimize the residual sum of squares between the observed targets in the dataset, and the targets predicted by the linear approximation. Linear regression is used for finding linear relationship between target and one or more predictors.

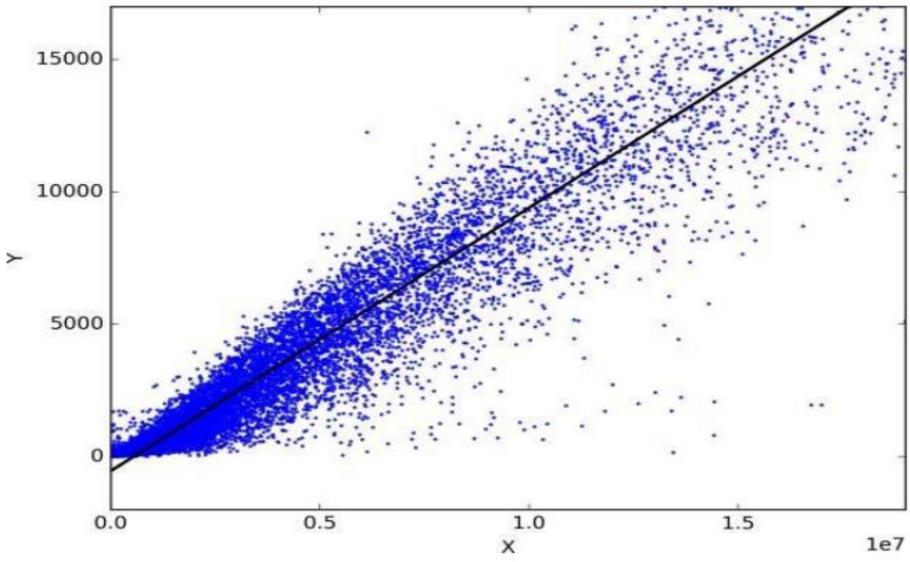

Figure 4.1: Linear Regression

- <u>Logistic Regression</u>**:** Logistic Regression is the appropriate regression analysis to conduct when the dependent variable is dichotomous (binary). The logistic model (or logit model) is used to model the probability of a certain classor event existing such as pass/fail, win/lose, alive/dead or healthy/sick. This can be extended to model several classes of events such as determining whether an image contains a cat, dog, lion, etc. Each object being detected in the image would be assigned a probability between 0 and 1 and the sum adding to one.

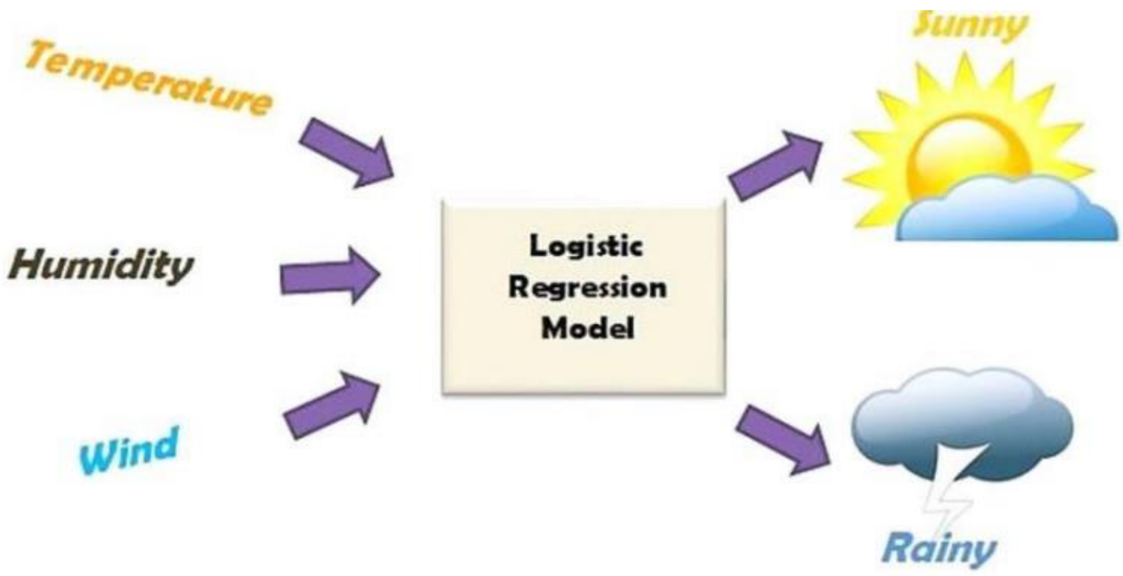

Figure 4.2: Logistic Regression

- <u>Decision Tree</u>**:** Decision Tree algorithm belongs to the family of supervised learning algorithms. Unlike other supervised learning algorithms, the decision tree algorithm can be used for solving regression and classification problems too. The goal of using a Decision Tree is to create a training model that can use to predict the class or value of the target variable by learning simple decision rules inferred from prior data(training data). In Decision Trees, for predicting a class label for a record we start from the root of the tree. We compare the values of the root attribute with the record‟s attribute. On the basis of comparison, we follow the branch corresponding to that value and jump to the next node.

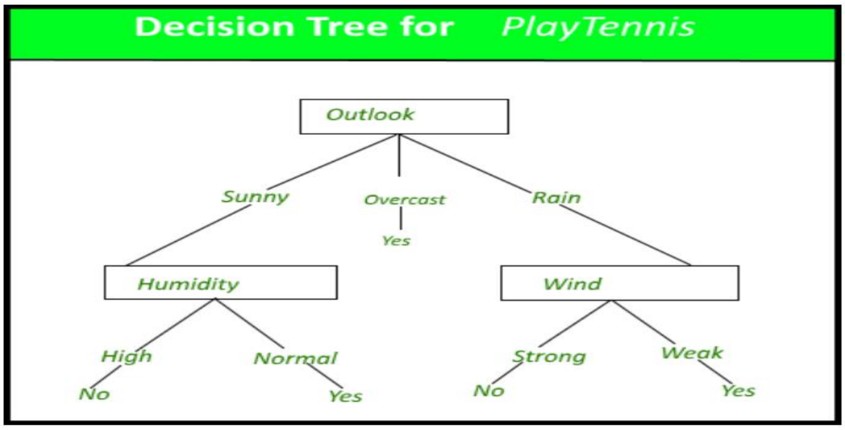

Figure 4.3: Decision Tree

The provided information appears to be the performance metrics of a decision tree classifier for a crop recommendation system. The table provides precision, recall, and F1-score for different crop classes such as apple, banana, blackgram, chickpea, coconut, coffee, cotton, grapes, jute, kidneybeans, lentil, maize, mango, mothbeans, mungbean, muskmelon, orange, papaya, pigeonpeas, pomegranate, and rice.

Finally, the accuracy, macro average, and weighted average are reported. The accuracy indicates the overall correctness of the classifier across all classes. The macro average and weighted average represent the average of precision, recall, and F1-score across all classes, with the weighted average considering the number of instances in each class.

```
Decision Tree Train Accuracy is : 0.9981
Decision Tree Test Accuracy is : 0.9909
```

	precision	recall	f1-score	support
apple	1.0000	1.0000	1.0000	28
banana	0.9600	1.0000	0.9796	24
blackgram	0.9615	1.0000	0.9804	25
chickpea	1.0000	1.0000	1.0000	37
coconut	1.0000	1.0000	1.0000	37
coffee	1.0000	0.9697	0.9846	33
cotton	1.0000	1.0000	1.0000	28
grapes	1.0000	1.0000	1.0000	37
jute	0.9259	1.0000	0.9615	25
kidneybeans	1.0000	1.0000	1.0000	33
lentil	1.0000	0.9677	0.9836	31
maize	0.9697	1.0000	0.9846	32
mango	1.0000	1.0000	1.0000	31
mothbeans	0.9630	0.9630	0.9630	27
mungbean	1.0000	1.0000	1.0000	24
muskmelon	1.0000	1.0000	1.0000	33
orange	1.0000	1.0000	1.0000	30
papaya	1.0000	0.9600	0.9796	25
pigeonpeas	1.0000	1.0000	1.0000	30
pomegranate	1.0000	1.0000	1.0000	27
rice	1.0000	0.9375	0.9677	32
...				
accuracy			0.9909	660
macro avg	0.9900	0.9908	0.9902	660
weighted avg	0.9913	0.9909	0.9909	660

Table 4.1: Performance Matrices of Decision Tree

- Gaussian Naive Bayes: Gaussian Naive Bayes is a variant of Naive Bayes that follows Gaussian normal distribution and supports continuous data. Naive Bayes methods are a set of supervised learning algorithms based on applying Bayes‟ theorem with the “naive” assumption of conditional independence between every pair of features given the value of the class variable.

$$P(x_i \mid y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right)$$

- <u>Support Vector Machine (SVM) algorithm</u>: Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.

  The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.

  SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine. Consider the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane:

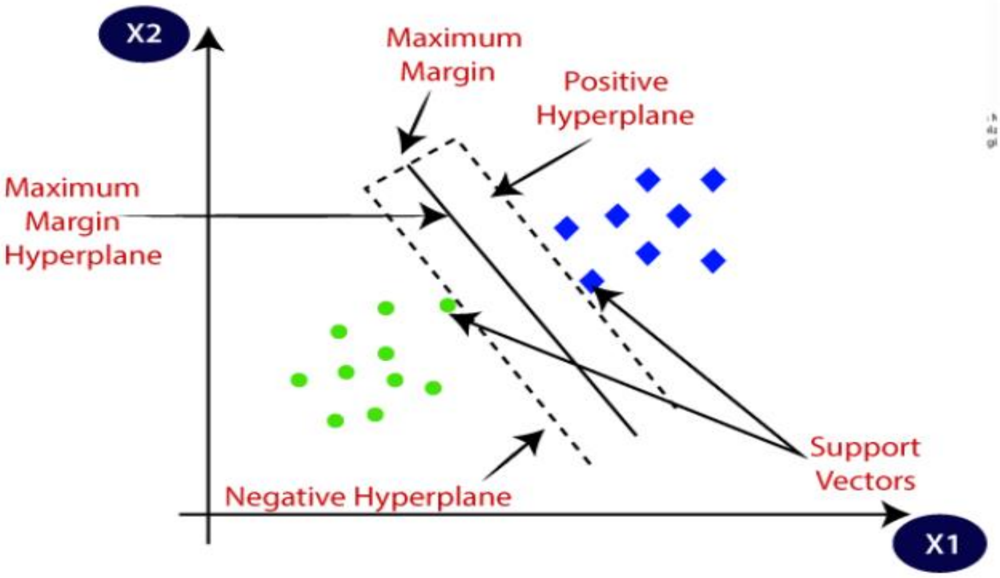

Figure 4.4: SVM Hyperplane

- <u>Random Forest:</u> A random forest is a machine learning technique that"s used to solve regression and classification problems. It utilizes ensemble learning, which is a technique that combines many classifiers to provide solutions to complex problems.

  A random forest algorithm consists of many decision trees. The „forest" generated by the random forest algorithm is trained through bagging or bootstrap aggregating.

Bagging is an ensemble meta-algorithm that improves the accuracy of machine learning algorithms.

The (random forest) algorithm establishes the outcome based on the predictions of the decision trees. It predicts by taking the average or mean of the output from various trees. Increasing the number of trees increases the precision of the outcome.

A random forest eradicates the limitations of a decision tree algorithm. It reduces the over fitting of datasets and increases precision. It generates predictions without requiring many configurations in packages (like scikit-learn)

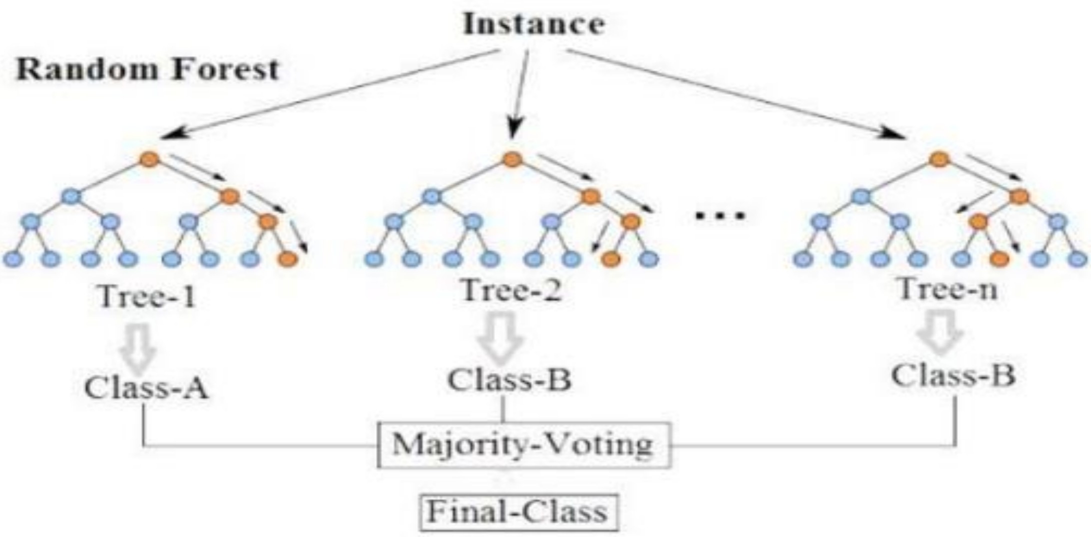

Figure 4.5: Random Forest

The provided information appears to be the performance metrics of a Random Forest classifier for a crop recommendation system. Here is a breakdown of the metrics:

- Train Accuracy: The accuracy of the Random Forest classifier on the training dataset is 0.9994, indicating that it correctly predicted the crop type with a high accuracy during the training phase.
- Test Accuracy: The accuracy of the Random Forest classifier on the test dataset is 0.9924, suggesting that it performed well in predicting the crop type on unseen data.

Precision, recall, and F1-score are evaluation metrics used to assess the performance of a classifier on each individual class. The metrics are often reported for each class separately, as well as their average values across all classes.

Random Forest Train Accuracy is : 0.9994
Random Forest Test Accuracy is : 0.9924

	precision	recall	f1-score	support
apple	1.0000	1.0000	1.0000	28
banana	1.0000	1.0000	1.0000	24
blackgram	0.9615	1.0000	0.9804	25
chickpea	1.0000	1.0000	1.0000	37
coconut	1.0000	1.0000	1.0000	37
coffee	1.0000	1.0000	1.0000	33
cotton	1.0000	1.0000	1.0000	28
grapes	1.0000	1.0000	1.0000	37
jute	0.8621	1.0000	0.9259	25
kidneybeans	1.0000	1.0000	1.0000	33
lentil	1.0000	1.0000	1.0000	31
maize	1.0000	1.0000	1.0000	32
mango	1.0000	1.0000	1.0000	31
mothbeans	1.0000	0.9630	0.9811	27
mungbean	1.0000	1.0000	1.0000	24
muskmelon	1.0000	1.0000	1.0000	33
orange	1.0000	1.0000	1.0000	30
papaya	1.0000	1.0000	1.0000	25
pigeonpeas	1.0000	1.0000	1.0000	30
pomegranate	1.0000	1.0000	1.0000	27
rice	1.0000	0.8750	0.9333	32
...				
accuracy			0.9924	660
macro avg	0.9920	0.9926	0.9919	660
weighted avg	0.9933	0.9924	0.9924	660

Table 4.2: Performance Matrices of Random Forest

- XGBoost: is a decision-tree-based ensemble Machine Learning algorithm that uses a gradient boosting framework. In prediction problems involving unstructured data (images, text, etc.) artificial neural networks tend to outperform all other algorithms or frameworks. However, when it comes to small-to-medium structured/tabular data, decision tree based algorithms are considered best-in-class right now.

- Confusion Matrix: A confusion matrix is a table that is used to define the performance of a classification algorithm. A confusion matrix visualizes and summarizes the performance of a classification algorithm.

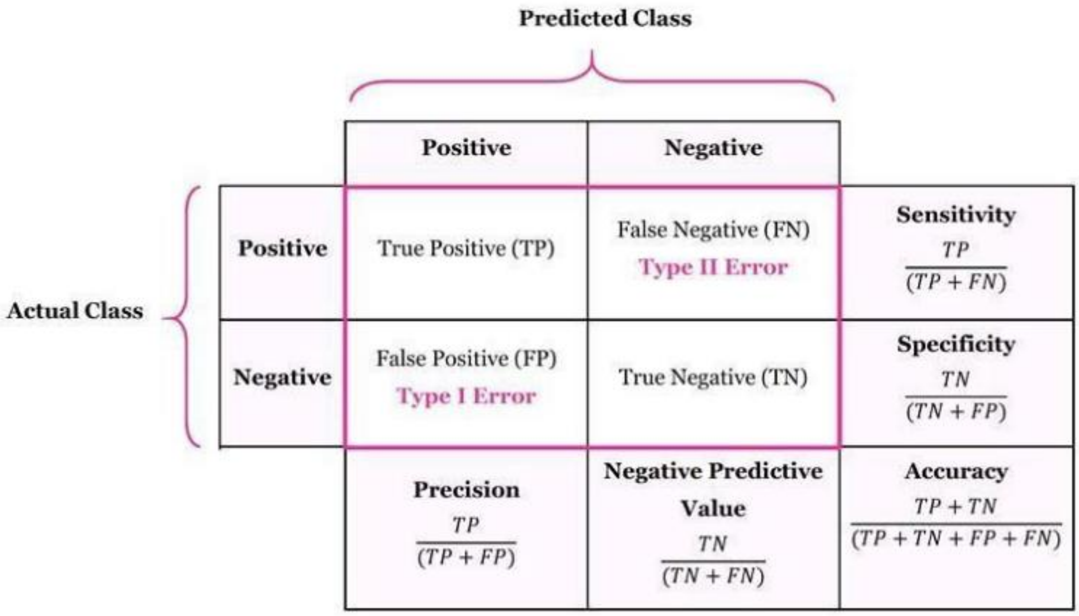

Figure 4.6: Confusion Matrix

The information provided in below table appears to be a confusion matrix showing the predicted versus actual values for a crop recommendation system. A confusion matrix is a table that visualizes the performance of a classification model by showing the counts of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions.

However, the formatting of the confusion matrix is not clear, making it difficult to interpret the results accurately. It seems that the matrix consists of predicted and actual values for multiple classes, but the exact layout and values are not discernible from the given information.

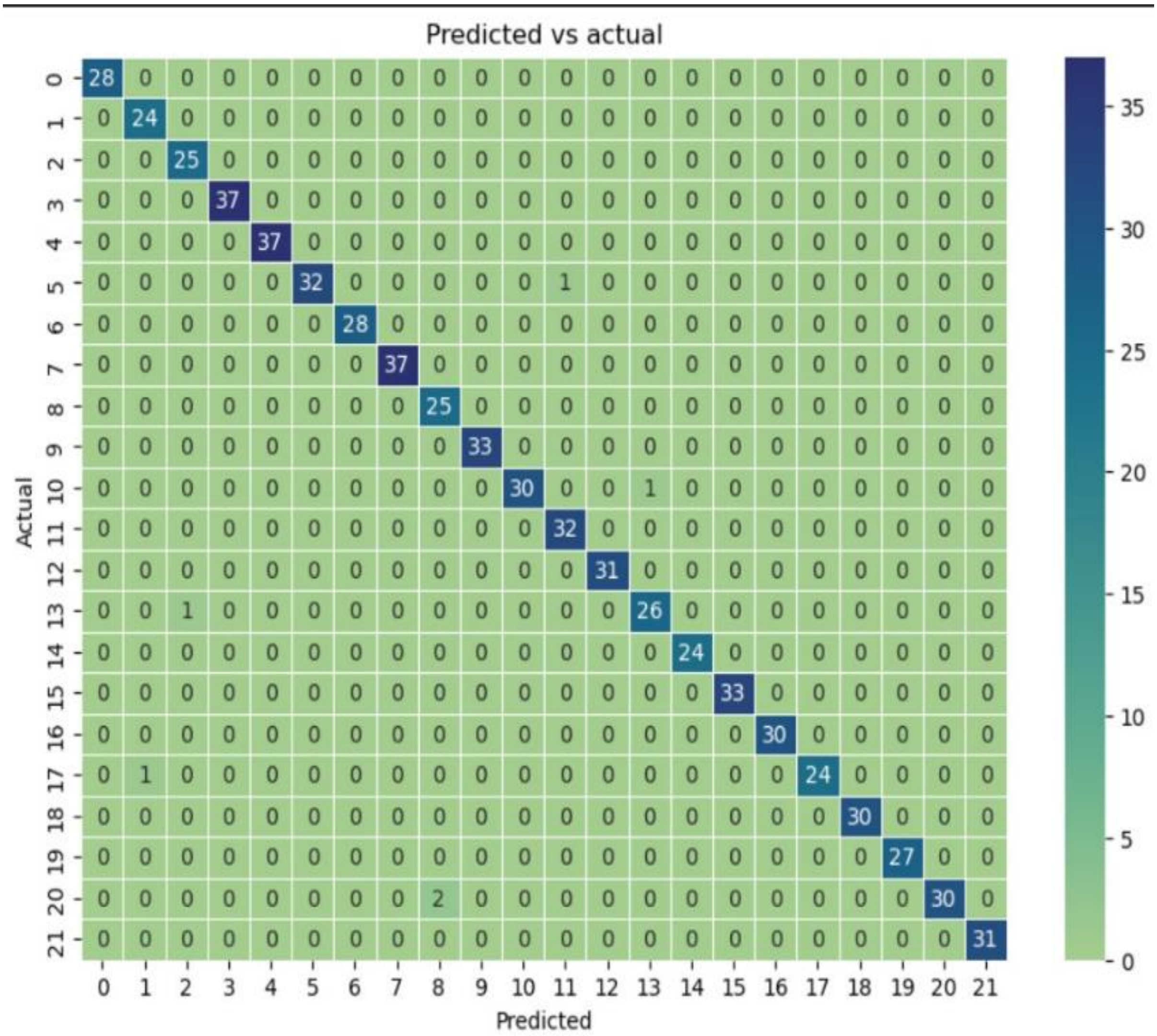

Table 4.3: Predicted Vs Actual

Chapter 5

EXPERIMENTAL RESULTS / OUTPUT

- **Home Page:**

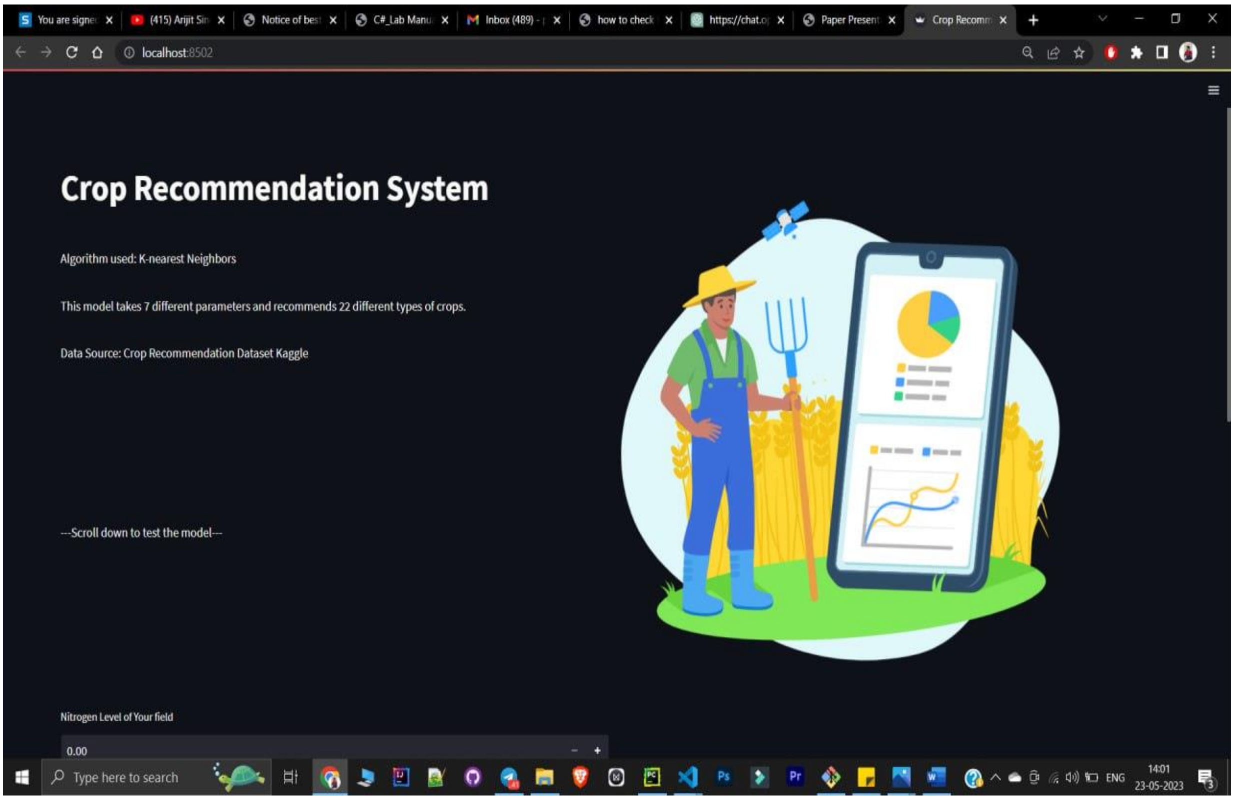

Figure 5.1: Home Page

Our web page is dedicated to a Crop Recommendation System that utilizes the powerful K-nearest Neighbors algorithm. This intelligent system aims to assist users in making informed decisions about crop selection based on their specific requirements and environmental factors. By inputting seven different parameters, such as soil type, climate conditions, and nutrient availability, users can receive personalized recommendations for 22 different types of crops.

The algorithm we employ, K-nearest Neighbors, analyzes the user's input and compares it to a comprehensive Crop Recommendation Dataset sourced from Kaggle. This dataset contains valuable information about various crops and their attributes. By leveraging the algorithm's ability to find similarities among data points, our system identifies the most suitable crops for the user's specific conditions.To make the system easily accessible and user-friendly, we have created a user interface on our web page. Users can scroll down to the interactive section and input their desired parameters to receive real-time crop recommendations. The system processes the inputs using the K-nearest Neighbors algorithm and presents the results promptly.

Our web page serves as a valuable resource for farmers, agriculture enthusiasts, and anyone interested in optimizing crop selection decisions. By harnessing the power of machine learning and providing personalized recommendations, we aim to enhance agricultural productivity and contribute to sustainable farming practices.

- **Recommended Crop :**

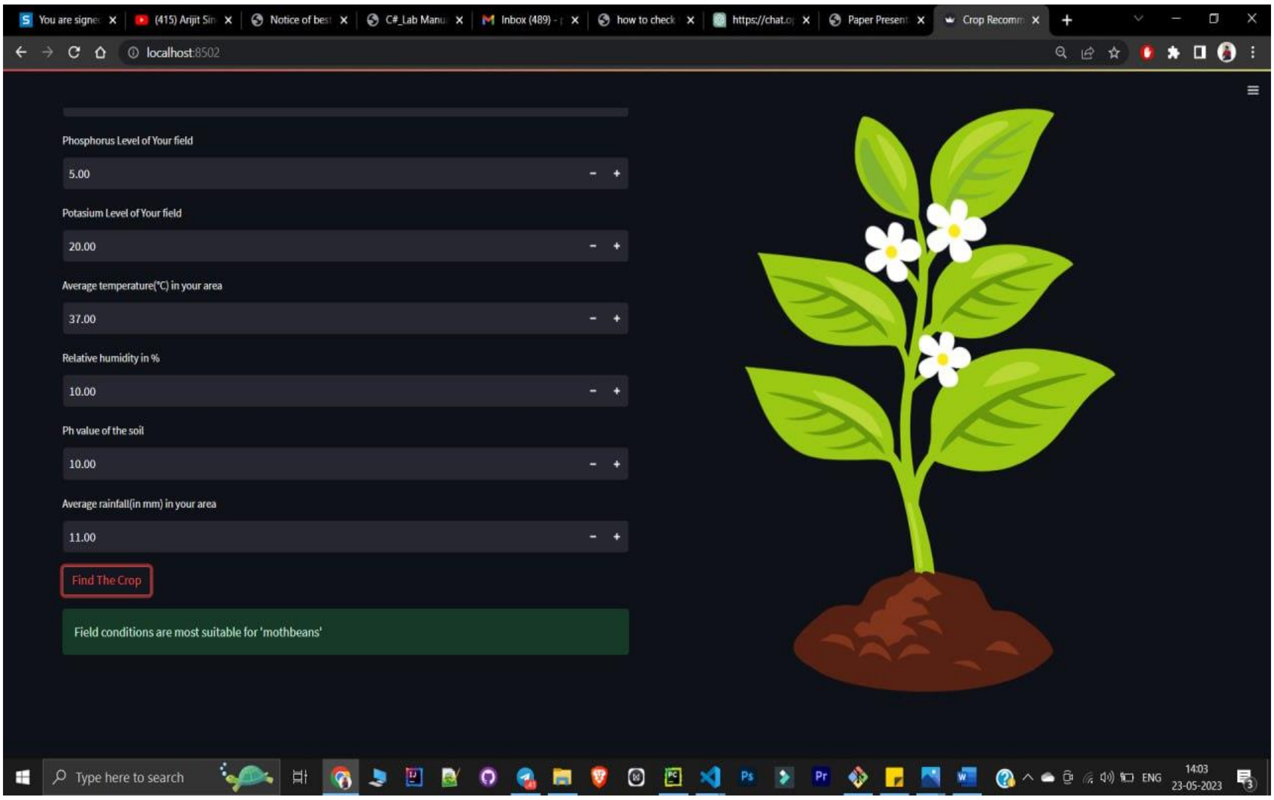

Figure 5.2: Crop Recommendation

Our web page is designed to assist farmers and agriculture enthusiasts in making informed decisions about crop selection for their fields. Through the use of advanced algorithms and comprehensive data analysis, our Crop Recommendation System provides personalized crop recommendations based on various factors such as soil conditions, climate, and environmental parameters.

By leveraging the power of machine learning and utilizing a vast dataset, our system takes into account crucial information about field conditions, including nutrient levels, temperature, humidity, pH value, and rainfall. These parameters are carefully analyzed to generate tailored crop recommendations that suit the specific needs and requirements of farmers. We are committed to supporting sustainable farming practices and empowering farmers with data-driven insights for improved decision-making. Our web page serves as a comprehensive resource for crop selection, farming guidance, and agricultural knowledge, ultimately helping farmers achieve better outcomes and drive the success of their farming endeavors.

Chapter 6

CONCLUSIONS AND FUTURE SCOPE

Conclusion:

In conclusion, the Crop Recommendation System offers a valuable tool for farmers and agriculture enthusiasts seeking guidance in selecting the most suitable crops for their fields. By leveraging advanced algorithms and comprehensive data analysis, the system provides personalized recommendations based on various field conditions, such as nutrient levels, temperature, humidity, pH value, and rainfall.

The web pages serves as a reliable source of information, offering not only crop recommendations but also detailed crop profiles and best practices for cultivation. It aims to empower farmers with data-driven insights to maximize agricultural productivity, optimize resource allocation, and ultimately contribute to the success and profitability of farming operations.

By incorporating machine learning and utilizing a vast dataset, the system ensures accurate and reliable recommendations tailored to the unique needs and requirements of farmers. It strives to support sustainable farming practices and help farmers make informed decisions for improved outcomes in their farming endeavors.

Overall, the Crop Recommendation System is a valuable resource that aims to enhance decision-making capabilities in crop selection, provide agricultural knowledge, and assist farmers in achieving better results in their farming practices.

Future Scope:

The Crop Recommendation System web page has a promising future with several potential areas of expansion and improvement. Some of the future scope possibilities include:

- Enhanced Data Integration: The system can benefit from integrating additional sources of data, such as satellite imagery, weather forecasts, and soil composition analysis, to further refine the crop recommendations. This can provide farmers with more accurate and comprehensive insights into their field conditions.

- Expansion of Crop Database: The system can be continuously updated with new crop varieties and their characteristics to offer a wider range of crop recommendations. Including more local and region-specific crops can cater to the diverse needs of farmers in different areas.

- Mobile Application Development: Developing a mobile application version of the Crop Recommendation System can provide farmers with on-the-go access to crop recommendations, field monitoring, and other essential features. Mobile apps can offer convenience and ease of use, making the system more accessible to a larger user base.

- Collaborations and Partnerships: Collaborating with agricultural research institutions, government agencies, and industry stakeholders can enhance the system's credibility and access to valuable data. Partnering with experts in the field can also lead to the development of innovative features and technologies within the system.

References

[1]Rakesh Kumar , M.P. Singh, Prabhat Kumar and J.P. Singh, “Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique”, International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials, 2015

[2] Haedong Lee and Aekyung Moon, “Development of Yield Prediction System Based on Real-time Agricultural Meteorological Information”, 16th International Conference on Advanced Communication Technology, 2014

[3] T.R. Lekhaa, “Efficient Crop Yield and Pesticide Prediction for Improving Agricultural Economy using Data Mining Techniques”, International Journal of Modern Trends in Engineering and Science (IJMTES), 2016, Volume 03, Issue 10

[4] Jay Gholap, Anurag Ingole, Jayesh Gohil, Shailesh Gargade and Vahida Attar, “Soil Data Analysis Using Classification Techniques and Soil Attribute Prediction”, International Journal of Computer Science Issues, Volume 9, Issue 3

[5] Anshal Savla, Parul Dhawan, Himtanaya Bhadada, Nivedita Israni, Alisha Mandholia, Sanya Bhardwaj (2015), „Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture‟, Innovations in Information, Embedded and Communication systems (ICIIECS). Intelligent Crop Recommendation System using ML Chapter 10

[6] AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms. Zeel Doshi , Subhash Nadkarni , Rashi Agrawal, Prof. Neepa Shah.

[7] Crop Recommendation System for Precision Agriculture S.Pudumalar*, E.Ramanujam*, R.Harine Rajashree, C.Kavya, T.Kiruthika, J.Nisha.

[8] Tom M. Mitchell, Machine Learning, India Edition 2013, McGraw Hill Educa- tion.

[9] https://data.gov.in/g [10] Kagglehttps://www.kaggle.com/notebook

Paper Publication Certificate's:

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN SCIENCE, COMMUNICATION AND TECHNOLOGY

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

IJARSCT

CERTIFICATE OF PUBLICATION

INTERNATIONAL STANDARD SERIAL NUMBER

ISSN NO: 2581-9429

THIS IS TO CERTIFY THAT

Tejal Survase

SVERI's College of Engineering, Pandharpur, Maharashtra, India

HAS PUBLISHED A RESEARCH PAPER ENTITLED

Intelligent Crop Recommendation System using Machine Learning

IN IJARSCT, VOLUME 3, ISSUE 11, MAY 2023

Certificate No: 052023-A3416

www.ijarsct.co.in

Crossref

DOI: 10.48175/IJARSCT-10585

www.doi.org

www.crossref.org

Scientific Journal Impact Factor 7.301 Journal Impact Factor

www.sjifactor.com

ISSN 2581-9429 IJARSCT

Editor-in-Chief


International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
IJARSCT
INTERNATIONAL STANDARD SERIAL NUMBER
ISSN NO: 2581-9429
CERTIFICATE OF PUBLICATION
THIS IS TO CERTIFY THAT
Siddhi Velapure
SVERI's College of Engineering, Pandharpur, Maharashtra, India
HAS PUBLISHED A RESEARCH PAPER ENTITLED
Intelligent Crop Recommendation System using Machine Learning
IN IJARSCT, VOLUME 3, ISSUE 11, MAY 2023
Certificate No: 052023-A3417
www.ijarsct.co.in
Crossref
DOI: 10.48175/IJARSCT-10585
www.doi.org
www.crossref.org
7.301
Journal Impact Factor
www.sjifactor.com
Editor-in-Chief

International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

IJARSCT

INTERNATIONAL STANDARD SERIAL NUMBER
ISSN NO: 2581-9429

Certificate of Publication

This is to certify that

Mayuri Swami

SVERI's College of Engineering, Pandharpur, Maharashtra, India

Has published a research paper entitled

Intelligent Crop Recommendation System using Machine Learning

in IJARSCT, Volume 3, Issue 11, May 2023

Certificate No: 052023-A3418
www.ijarsct.co.in

Crossref
DOI: 10.48175/IJARSCT-10585
www.doi.org
www.crossref.org

Scientific Journal Impact Factor 7.301 Journal Impact Factor
www.sjifactor.com

Editor-in-Chief

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN SCIENCE, COMMUNICATION AND TECHNOLOGY

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

IJARSCT

CERTIFICATE OF PUBLICATION

INTERNATIONAL STANDARD SERIAL NUMBER
ISSN NO: 2581-9429

THIS IS TO CERTIFY THAT

Ratan Kore

SVERI's College of Engineering, Pandharpur, Maharashtra, India

HAS PUBLISHED A RESEARCH PAPER ENTITLED

Intelligent Crop Recommendation System using Machine Learning

IN IJARSCT, VOLUME 3, ISSUE 11, MAY 2023

Certificate No: 052023-A3419
www.ijarsct.co.in

Crossref
DOI: 10.48175/IJARSCT-10585
www.doi.org
www.crossref.org

Scientific Journal Impact Factor 7.301
www.sjifactor.com

Editor-in-Chief

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN SCIENCE, COMMUNICATION AND TECHNOLOGY

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

IJARSCT

CERTIFICATE OF PUBLICATION

INTERNATIONAL STANDARD SERIAL NUMBER
ISSN NO: 2581-9429

THIS IS TO CERTIFY THAT

Prapti Atkale

SVERI's College of Engineering, Pandharpur, Maharashtra, India

HAS PUBLISHED A RESEARCH PAPER ENTITLED

Intelligent Crop Recommendation System using Machine Learning

IN IJARSCT, VOLUME 3, ISSUE 11, MAY 2023

Certificate No: 052023-A3420
www.ijarsct.co.in

Crossref
DOI: 10.48175/IJARSCT-10585
www.doi.org
www.crossref.org

7.301
Scientific Journal Impact Factor
www.sjifactor.com

Editor-in-Chief

Intelligent Crop Recommendation System using Machine Learning

Sanjivani Kadam[1], Tejal Survase[2], Siddhi Velapure[3], Mayuri Swami[4], Ratan Kore[5], Prapti Atkale[6]

Assistant Professor, Department of Computer Science and Engineering[1] Students, Department of Computer Science and Engineering[2,3,4,5,6] SVERI's College of Engineering, Pandharpur, Maharashtra, India

Punyashlok Ahilyadevi Holkar Solapur University, Gopalpur, Pandharpur, Maharashtra, India

Abstract: *Intelligent crop recommendation systems have gained significant attention in modern agriculture for their potential to optimize crop selection and enhance agricultural productivity. This report aims to provide a comprehensive analysis of existing intelligent crop recommendation systems and propose a novel framework for future system development. The report explores the current state-of-the-art in the field, identifies the key components and functionalities of existing systems, and evaluates their strengths and limitations. Building upon this analysis, a proposed system framework is presented, encompassing data acquisition, preprocessing, machine learning algorithms, recommendation generation, and user interface. The proposed system addresses the limitations of existing systems and leverages emerging technologies for improved accuracy, scalability, and sustainability. The report concludes with a discussion on the potential impact of the proposed system on agricultural practices and highlights future research directions*

Keywords: *Agriculture, Maximum Crop Yield, Fertilizer Suggestion, Environmental Factor, Economic Factor, Machine Learning(ML), Plant Disease Classification*

I. INTRODUCTION

In recent years, there has been a growing need for efficient and sustainable agricultural practices to meet the increasing demand for food production. One critical aspect of successful farming is the selection of suitable crops for specific environmental conditions, soil characteristics, and market demands. However, this task can be challenging for farmers due to the complexity and variability of these factors. To address this challenge, intelligent crop recommendation systems have emerged as valuable tools for optimizing crop selection. These systems leverage advanced technologies such as machine learning, data analytics, and artificial intelligence to provide data-driven recommendations to farmers. By analyzing large volumes of agricultural data, including historical crop yields, climate data, soil information, and market trends, these systems can offer personalized and accurate crop recommendations. The fundamental goal of an intelligent crop recommendation system is to enhance agricultural productivity and profitability while minimizing environmental impact. By suggesting the most suitable crops based on the specific conditions of a farm, these systems can help farmers make informed decisions and maximize their yields. Additionally, intelligent crop recommendation systems can aid in resource management by optimizing the use of fertilizers, pesticides, and water, thus reducing waste and potential harm to the environment.

The success of intelligent crop recommendation systems relies heavily on the quality and diversity of the data used for analysis. Integration of various data sources, such as satellite imagery, weather stations, soil sensors, and market data, provides a comprehensive understanding of the agricultural landscape. With access to real-time and historical data, these systems can continuously learn and adapt, improving the accuracy of their recommendations over time.

Furthermore, intelligent crop recommendation systems can consider multiple objectives and constraints, such as crop rotation, disease resistance, and market demand, to provide holistic and tailored recommendations to farmers. By considering a wide range of factors, these systems help optimize crop selection for long-term sustainability and economic viability.

II.LITERATURE REVIEW

[1] S.Pudumalar*, E.Ramanujam*, R.HarineRajashreeń, C.Kavyań, T.Kiruthikań, J.Nishań. :: In a country like India in which agriculture plays a dominant role. The prosperity of nation increases because of farmers prosperity. Thus our work would help farmers in sowing the right seed based on soil requirements to enhance productivityand gain profit outof such

a technique. Thus the farmers can plant the right cropimproving their income and also increasing the gross productivity of the state. Our future work is focused on an upgraded data set with a large number of attributes and also implements yield prediction

[2] Shikha Ujjainia, Pratima Gautam, S. Veenadhari :: From the above analysis, it has been concluded that technology has achieved that level of competency by which agriculture can easily predict cropyielding production. Machine learning technology proposed the model integrated form of a concept. By estimating the different parameters of biosystems has been recognized that the technology used for making the crop yielding prediction device is very much variegated. The parameters of biosystems vary concerning changing location and a single concept of algorithm is not sufficient to fulfill the requirement of crop prediction.

[3] Mahendra Choudhary, Rohit Sartandel, Anish Arun, LeenaLadge SIES Graduate School of Technology, Maharashtra, India Corresponding author: Rohit Sartandel, Email:rohit.sartandel18@siesgst.ac.i :: At present, our farmers are not using technology and analytics productively, so there is a probability of fallacious crop selection for cultivation, which will reduce theirincome. To fend off such dropping, we have developed a farmer-friendlysystem witha graphical user interface (GUI) that will predict which crop would be the best fit for a specific plot of land. This system will also furnish details on the essential nutrients to add to the soil and help them to pick out crop diseases. As a result, farmers are more likely to make the proper decisions for crop selection, and farmers will benefit from earning more profits. With the help of more composite algorithms, the efficiency of the model can be increased. To boost the efficiency of crop selection and yield, incorporate the crop recommendation system with yield prediction. By increasing the dataset's size and by including more crop information and pictures of infected plants, the efficiency of the model can be increased.

III.PROPOSED SYSTEM

The proposed system harnesses the power of data analysis technology to actively monitor and update crop yield rates.Its primary objective is to introduce a robust crop selection method that effectively tackles the challenges faced by farmers and the agricultural sector. By maximizing crop yield rates, this approach seeks to make a substantial contribution to the overall improvement of the Indian economy.

One of the key strengths of the system lies in its ability to consider diverse land conditions and accurately identify the quality of crops. Through this process, the system distinguishes high-quality crops, enabling farmers to make informed decisions about their cultivation choices. To enhance the precision of crop predictions, the system leverages anensemble of classifiers, such as Decision Tree and Random Forest classifiers. By combining the outputs of multiple classifiers, the system enhances decision-making capabilities, leading to more reliable predictions and improved efficiency.

The integration of ensemble classifiers into the decision-making process further strengthens the overall system performance. By incorporating the insights and predictions from multiple classifiers, the system attains higher accuracy and better adaptability. This integration ensures that the proposed system maximizes the potential of data analysis technology and effectively addresses the challenges faced by farmers and the agricultural sector. In summary, the proposed system capitalizes on data analysis technology to continually monitor and update crop yield rates. Its core objective is to introduce a comprehensive crop selection method that overcomes existing agricultural challenges. By prioritizing the maximization of crop yield rates, the system aims to contribute significantly to the advancement of the Indian economy. Through its consideration of diverse land conditions, accurate crop quality identification, and integration of ensemble classifiers, the system enhances prediction accuracy, decision-making capabilities, and overall system efficiency.

2.1 Being System

The proposed project introduces a novel model aimed at addressing existing agricultural challenges. The system's key focus is to guide farmers in maximizing crop yield and selecting the most profitable crops for specific regions.

By considering economic and environmental conditions, the proposed model offers crop selection recommendations that prioritize both maximizing yield and meeting the country's increasing food demand. The model predicts crop yield based on factors such as state, district, area, and season, while also assisting in determining the optimal timing for fertilizer application.

For production-related inputs, the user provides information on the state, district, season, crop, and area. For crop recommendation, the user provides details on the state, district, season, and area. Based on these inputs, the model predicts the crop yield for a specific crop and suggests the most profitable crop option, along with the ideal time for fertilizer application.

The primary objective of the system is to offer a diverse range of crops suitable for each season. By providing guidance in crop selection, the proposed system aims to minimize farmers' challenges and maximize overall yield.

Overall, the proposed system presents a valuable solution that assists farmers in selecting crops, maximizing yield, and overcoming difficulties in decision-making.

III.PROPOSED WORKFLOW

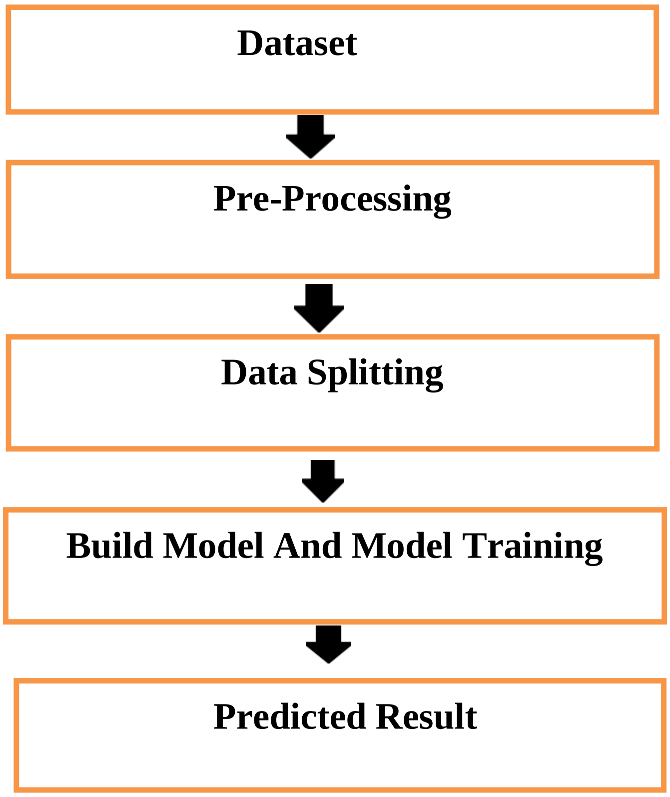

Fig: Proposed Workflow

IV.IMPLEMENTATION

4.1Input Dataset

The dataset provided to the system as inputs which include parameters like soil pH, humidity, N, P, K, temperature rainfall and the labels of the crops which are going to be given by a system as output.

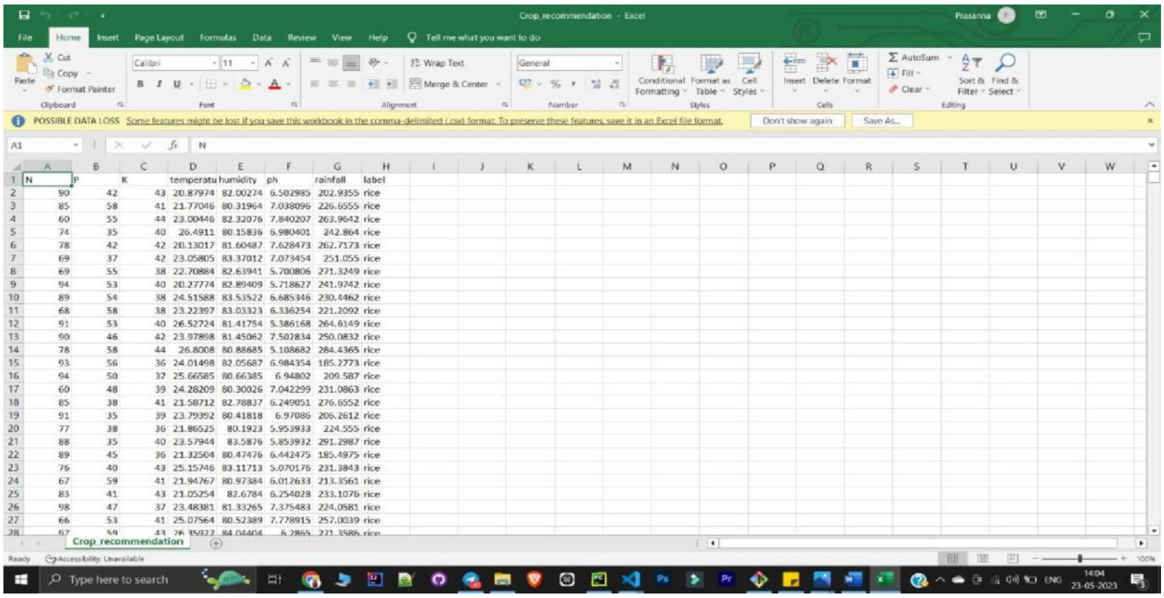

Fig:Input dataset

Fig : Dataset connector

Code Generation：

The generation of source code by using different libraries such as stremlit, Numpy, Image, pickle, requests, st_lottie.Here we connect the above dataset and write the programe which display frontend of the web.

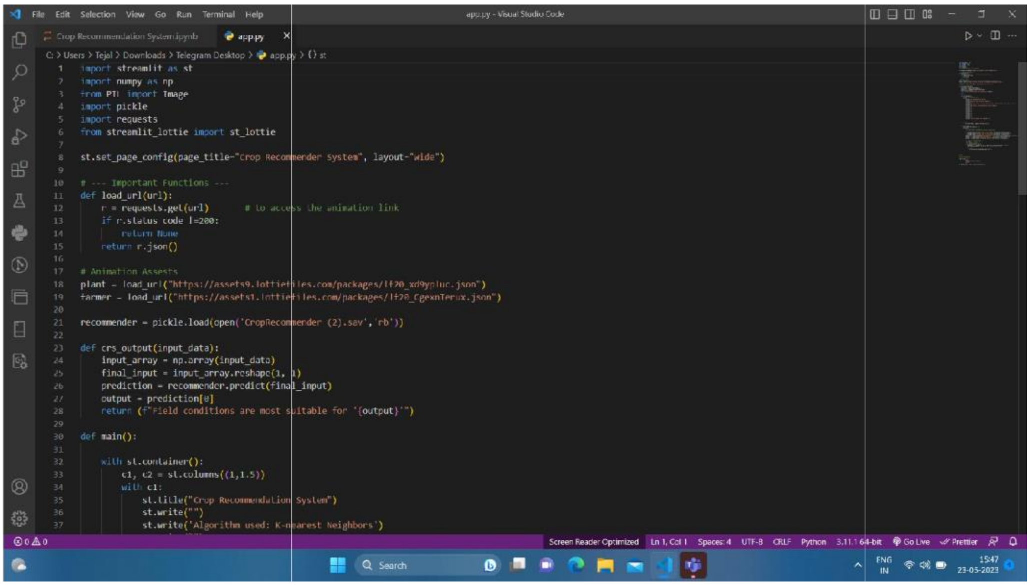

Fig: Source code

Web Front View :

After running source code we will get webpage as :

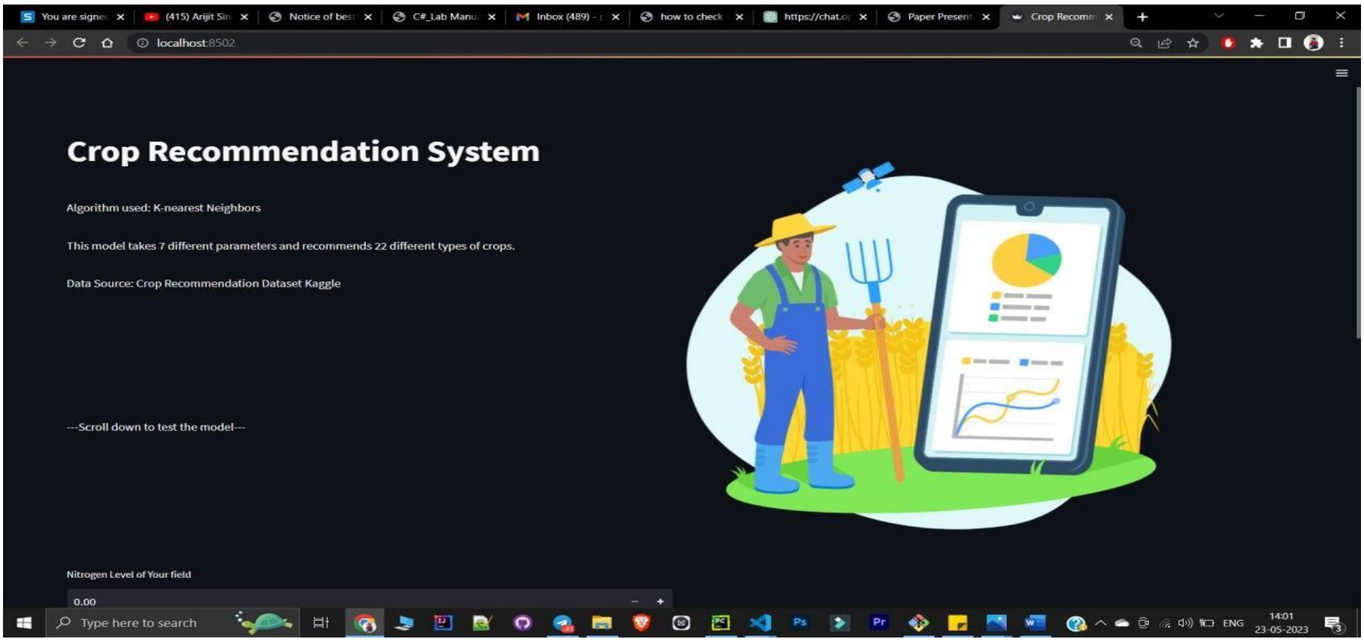

Fig : WebView

Final Output :

It shows final output that is the recommended crop .

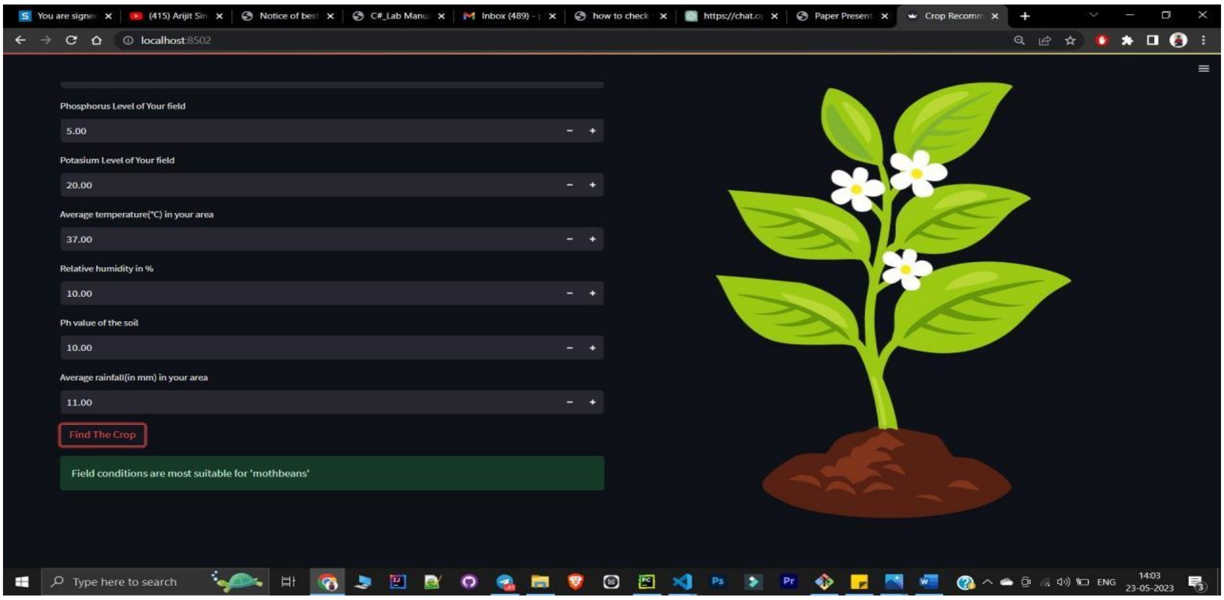

fig: Final output

V.TOOLS AND TECHNOLOGIES USED

- Python: Python is an abundance of powerful tools ready for scientific computing Packages. Thepackages like NumPy, Pandas and SciPy, Streamlite-lottie, Pillow. Pickle_mixing.
- NUMPY: NumPy is a powerful Python library which provides a fundamental infrastructure for numerical computing. It stands for "Numerical Python" and provides efficient data structures, mathematical functions, and tools for working with large arrays and matrices. NumPy forms the foundation for many scientific anddata analysis libraries in Python.
- SCIKIT-LEARN : It features numerous classification, clustering and regression algorithms like random forests, SVM and it furthermore supports Python scientific and numerical libraries like SciPy and NumPy.
- TENSORFLOW: In the TensorFlow [22] has an open source software library for numerical computation using data flow graphs.
- IDE used: Jupyter Notebook, PyCharm
- Dataset used: Kaggle Crop Dataset

VI.FUTURE SCOPE The

future scope of intelligent crop recommendation systems includes:

- Integration of IoT and sensor technologies for real-time data collection.
- Harnessing big data and advanced analytics techniques for improved recommendations.
- Utilizing remote sensing and satellite imagery for precise and proactive decision-making.
- Incorporating crop genetics and genomics to recommend resilient and high-yielding crop varieties.
- Developing user-friendly mobile applications and decision support tools for accessibility.
- Promoting collaboration and knowledge sharing among farmers and experts.

These advancements aim to enhance the accuracy, efficiency, and sustainability of crop recommendations, contributing to improved agricultural practices and global food security

VII.CONCLUSION

The proposed system framework for an intelligent crop recommendation system harnesses emerging technologies, and advanced machine learning algorithms, including ensemble methods and deep learning, enhance the accuracy and robustness of the recommendations by capturing complex patterns and relationships in the data. the proposed system framework represents a significant step forward in intelligent crop recommendation systems, offering a comprehensive and technologically advanced approach to assist farmers in making optimal crop decisions. By leveraging the power of data, analytics, and user-friendly interfaces, the system has the potential to revolutionize agriculture and drive sustainable farming practices in the future.

REFERENCES

[1]. Shilpa Mangesh Pande, Prem Kumar Ramesh, Anmol, B R Aishwarya, Karuna Rohilla and Kumar Shaurya, Crop Recommendation using Machine Learning Approach, Proceedings of the Fifth International Conference on Computing Methodologies and Communication (ICCMC 2021).

[2].https://www.researchgate.net/publication/331426761_Crop_Recommendation_System_to_Maximize_Crop_Yield_in_Ramtek_region_using_Machine_Learning

[3]. Haedong Lee and Aekyung Moon, “Development of Yield Prediction System Based on Real-time Agricultural Meteorological Information”, 16th International Conference on Advanced Communication Technology, 2014

[4]. Pudumalar, S., E. Ramanujam, R. Harine Rajashree, C. Kavya, T. Kiruthika, and J. Nisha. "Crop recommendation system for precision agriculture." In 2016 Eighth International Conference on Advanced Computing (ICoAC), pp. 32-36. IEEE, 2017.

[5]. T.R. Lekhaa, “Efficient Crop Yield and Pesticide Prediction forImproving Agricultural Economy using Data Mining Techniques”,International Journal of Modern Trends in Engineering and Science(IJMTES), 2016, Volume 03, Issue 10

[6].https://www.researchgate.net/publication/335799151_Crop_Recommendation_System_for_Precision_Agriculture