Project Title:

"Crop and Fertilizer Recommendation System"

Submitted in partial fulfillment of the requirements of the Course on

Artificial Intelligence (BTCS-T-PC-501)

of

Bachelor of Technology (B-Tech)
Biju Patnaik University of Technology
Department of
Computer Science and Engineering

Gandhi Institute for Technology (GIFT) Autonomous, Bhubaneswar

NAME OF THE STUDENT	REGISTRATION NUMBER	SEMESTER	BRANCH
Mohanty Hitesh Rabindranath	2201298111	5th	CSE-Genius
Syed Sadiqu Hussain	2201298227	5th	CSE-Genius

Asst. Prof. Aliva Pattaniak
(Name of the Guide)

Dr. Sujit Kumar Panda
(Head of the Department)

DECLARATION

I Mohanty Hitesh Rabindranath and Syed Sadiqu Hussain, hereby declare that this project report entitled **“Crop and Fertilizer Recommendation System”** for partial fulfillment of the requirements for the Degree of Bachelor of Technology is a Bonafede record of work carried out by me and the content included in this report is true to the best of my knowledge.

Date: 30/11/2024

Mohanty Hitesh Rabindranath
(2201298111)
Syed Sadiqu Hussain
(2201298227)

CERTIFICATE

This is to certify that the Subject-Specific-Project entitled **"Crop and Fertilizer Recommendation System"** has been carried out by Mohanty Hitesh Rabindranath (2201298111) and Syed Sadiqu Hussain (2201298227) and completed under my guidance and project meets the academic requirement of the subject Artificial Intelligence for Computer Science (BTBS-T-PC-501).

Signature of the Guide

Table of Contents

Abstract

In the modern agricultural landscape, data-driven solutions are pivotal in addressing challenges related to crop and fertilizer recommendations, optimizing productivity, and sustainable farming practices. This project, titled **"Intelligent Agricultural Decision Support System,"** aims to assist farmers by providing accurate and efficient recommendations for crop selection, fertilizer usage, and statistical insights based on soil and environmental data.

The system comprises three primary modules: a **Crop Recommendation System**, a **Fertilizer Recommendation System**, and a **Statistical Analysis Module**. Using advanced data processing techniques, the Crop Recommendation System suggests suitable crops based on soil and environmental parameters. Similarly, the Fertilizer Recommendation System advises appropriate fertilizers tailored to specific soil and crop needs. The Statistical Analysis Module provides farmers with graphical insights into historical trends, enabling informed decision-making.

The project incorporates a centralized database that stores and processes soil, crop, and fertilizer-related data, ensuring seamless integration across all modules. With a focus on user-friendly interactions, the system empowers farmers by leveraging technology to make sustainable and effective agricultural decisions.

This report discusses the system's design, implementation, and outcomes, highlighting its potential to transform traditional agricultural practices into data-driven methodologies. The project also underscores the importance of integrating predictive algorithms and statistical analysis to enhance the precision and efficiency of farming activities.

Keywords: Crop Recommendation, Fertilizer Recommendation, Statistical Analysis, Data-Driven Agriculture, Soil Data, Sustainable Farming, Agricultural Decision Support System, Predictive Analytics, Graphical Insights, Centralized Database

Introduction

Agriculture plays a pivotal role in the economy, serving as the backbone of many nations. However, with the increasing challenges posed by climate change, soil degradation, and inefficient farming practices, there is an urgent need to adopt technology-driven solutions to enhance agricultural productivity and sustainability. Traditional farming methods often rely on experience-based decisions, which may not always yield optimal results. This creates a significant opportunity for leveraging data-driven solutions to support informed decision-making in agriculture.

The **Crop and Fertilizer Recommendation System** is designed to empower farmers by integrating modern technologies like machine learning, predictive analytics, and centralized data storage. The system assists farmers in choosing the most suitable crops based on soil and environmental conditions, recommending fertilizers for better yield, and providing statistical insights to improve farming strategies. By centralizing and analyzing soil and crop data, this system aims to bridge the gap between traditional farming methods and modern technological advancements.

The project focuses on three core functionalities:

1. **Crop Recommendation**: Suggests optimal crops based on soil and environmental data using predictive models.
2. **Fertilizer Recommendation**: Provides recommendations for fertilizers tailored to the specific needs of the soil and crop.
3. **Statistical Analysis**: Offers graphical insights and analytics to help farmers understand trends and improve decision-making.

This system is not only a tool to enhance productivity but also a step toward sustainable agriculture. By making informed decisions, farmers can reduce waste, optimize resource utilization, and contribute to environmental conservation. Through the integration of advanced data analysis and machine learning techniques, the system provides a holistic solution to modern-day agricultural challenges.

.

Literature Survey

The development of intelligent agricultural systems has been extensively explored in recent years, with significant contributions in the domains of crop recommendation, fertilizer optimization, and data-driven analytics.

1. **Patel et al. (2020**) proposed a crop recommendation system leveraging machine learning algorithms such as Random Forest and SVM. Their model demonstrated high accuracy in predicting suitable crops based on soil attributes like pH, nitrogen, and potassium content.
2. **Sharma et al. (2019)** focused on fertilizer recommendations using decision tree models, addressing the issue of over-fertilization and its environmental impact. Their work highlighted the role of personalized fertilizer plans based on soil and crop requirements.
3. **Kumar et al. (2021)** developed a comprehensive agricultural decision support system integrating crop and fertilizer recommendations with weather prediction. Their ensemble-based approach significantly improved the accuracy of predictions.
4. **Jain and Singh (2022)** emphasized the importance of statistical analysis in agriculture. They designed a system that visualizes soil fertility trends and crop yield patterns, enabling better long-term planning for farmers.
5. **Chauhan et al. (2020)** utilized artificial neural networks (ANNs) to classify soil types and predict crop suitability. Their work showcased the efficiency of deep learning models in handling large agricultural datasets.

These studies form the foundation of this project, combining proven techniques like machine learning, predictive analytics, and statistical insights to create an integrated solution for modern agriculture. The proposed system builds upon these advancements, offering a holistic approach to assist farmers in improving crop yield and sustainability.

Problem Statement

Farmers often face challenges in selecting the most suitable crops and fertilizers for their fields due to limited access to accurate soil and environmental data. Additionally, the lack of analytical insights hampers effective decision-making, leading to reduced crop yield and sustainability. This project aims to address these challenges by developing an integrated system that recommends crops and fertilizers based on soil and environmental conditions, while also providing statistical insights to support informed agricultural practices.

Methodology

The project follows a structured approach to build a comprehensive system for crop and fertilizer recommendations:

1. **Data Collection**:
   - **Source**: Crop recommendation data was sourced from Kaggle and agricultural websites such as ICRISAT and AgroMet. Fertilizer data was based on a simulated dataset containing valid nutrient ranges derived from domain knowledge.
   - **Features**: The key features collected includes:

N	Numeric	Value of nitrogen present in the soil
P	Numeric	Value of phosphorus in soil
K	Numeric	Potassium content in soil
Humidity	Numeric	Humidity in %
Temperature	Numeric	Temperature in degree celsius
pH	Numeric	Soil quality which defines whether soil is acidic, basic or not
Rainfall	Numeric	Rainfall in mm
Label	Nominal	Name of the crop

2. **Exploratory Data Analysis (EDA)**:
   - Visualized trends in the data to understand relationships between variables.
   - Addressed missing values, handled outliers, and scaled features for uniformity.

3. **Model Selection and Training**:
   - **Model Comparison**: Evaluated multiple Machine learning models, including Logistic Regression (LR), Decision Trees, Random Forest (RF), and Support Vector Classifier (SVC).
   - **Training Process**:
     - Split data into 80% training and 20% testing sets for model validation.
     - Trained both crop and fertilizer recommendation models to ensure optimal accuracy.

4. **Model Testing and Evaluation**:
   - Finalized models based on performance metrics:
     - **Crop Recommendation**: Logistic Regression achieved the best results.
     - **Fertilizer Recommendation**: Random Forest performed optimally for predicting fertilizer needs.

5. **Model Deployment**:
   - **Platform**: The system was deployed as a web application using Streamlit for user accessibility.

- **Features**:
  - A user-friendly interface for farmers to input soil and environmental data.
  - Real-time crop and fertilizer recommendations.
  - Graphical insights for data visualization.

This methodology ensured that the models were not only accurate but also practical and accessible for end-users.

Source Code

GitHub Repository Link: *FarmSathi: Crop & Fertilizer Recommendation System*

The implementation of the project relies on the following libraries, frameworks, and APIs:

1. **NumPy**: Used for numerical computations and array manipulations.
2. **Pandas**: Essential for data preprocessing, manipulation, and analysis.
3. **Matplotlib & Seaborn**: Visualization libraries are used to create insightful plots and understand data relationships.
4. **Scikit-learn**: Utilized for training machine learning models, feature scaling, and model evaluation.
5. **Streamlit**: A Python framework for deploying the web application, providing a user-friendly interface for interaction.
6. **Pickle**: Used to serialize and save trained machine learning models for deployment.
7. **Microsoft Azure Document Intelligence API**:
   - Leveraged Optical Character Recognition (OCR) to extract textual data from uploaded soil reports or user-provided documents.
   - Enables automated preprocessing of textual information, enhancing usability and reducing manual data entry.
8. **Data Sources**:
   - Data sourced from Kaggle and agricultural websites like ICRISAT and AgroMet to simulate real-world use cases.

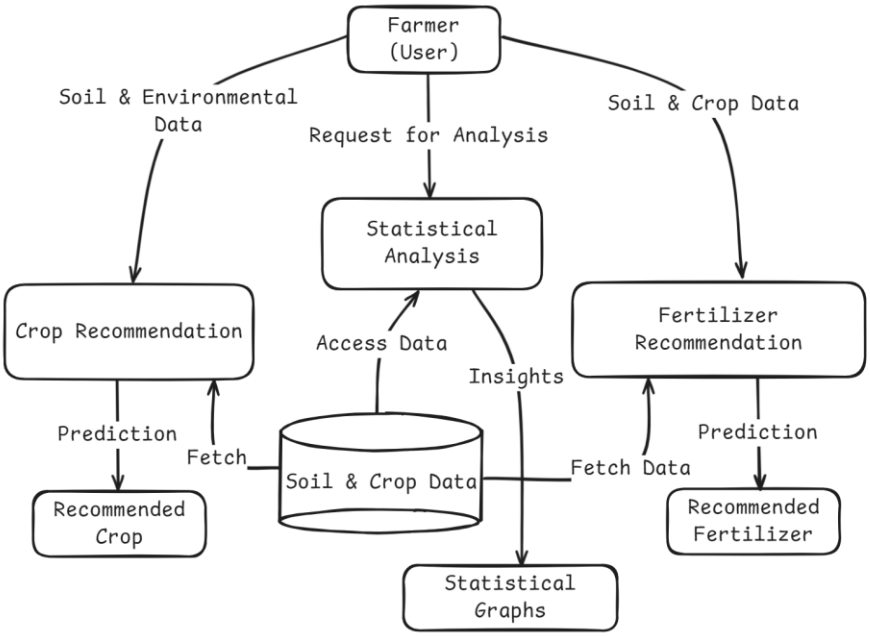
These libraries and APIs ensure efficient data handling, robust model training, seamless application deployment, and intelligent text extraction for user-contributed data.

System Design

The system is designed to assist farmers by providing crop and fertilizer recommendations based on soil, environmental, and crop data. The design consists of three major components:

1. **Crop Recommendation Module**
2. **Fertilizer Recommendation Module**
3. **Statistical Analysis Module**

Entire Workflow Diagram:



1. **User Input:**
   Farmers provide inputs such as soil type, environmental parameters, and crop data.
2. **Statistical Analysis Module:**
   This module processes the data to generate insights, such as nutrient deficiencies and historical crop performance trends. It also outputs **Statistical Graphs** to help users understand the data better.
3. **Crop Recommendation Module:**
   Based on soil and environmental data, this module fetches the necessary information from the database and predicts the most suitable crop using machine learning models.
4. **Fertilizer Recommendation Module:**
   Using soil and crop data, this module recommends fertilizers that optimize crop yield.

Result Analysis

1. Performance Metrics

The performance of the models was measured using Precision, Recall, F1-Score, and Support. The table below summarizes these results:

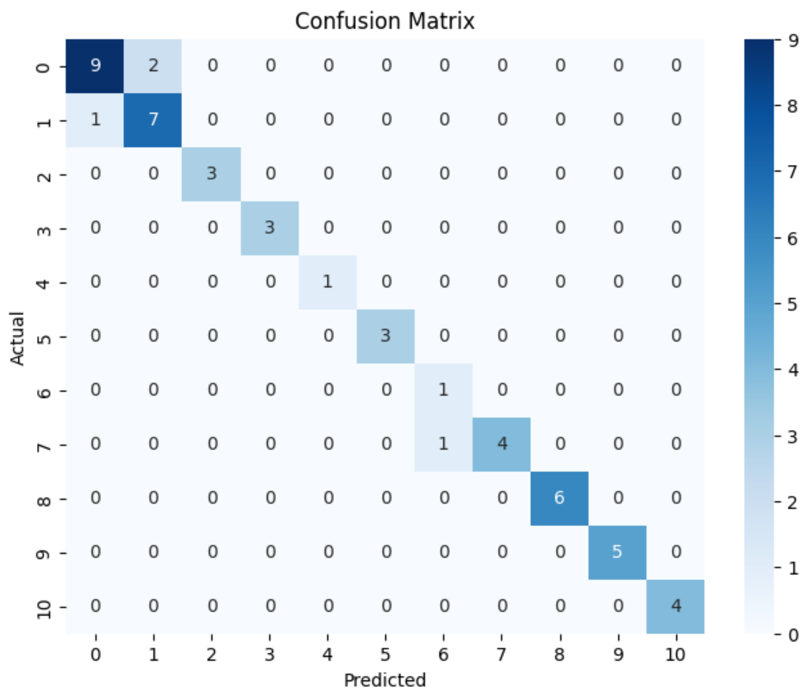
Model	Precision	Recall	F1-Score	Support
Crop	0.96	0.96	0.96	440
Fertilizer	0.93	0.95	0.93	50

- **Crop Recommendation**: Logistic Regression achieved an impressive **96% accuracy**, indicating its strong ability to classify crops based on soil and environmental features.
- **Fertilizer Recommendation**: Random Forest achieved a respectable **92% accuracy**, effectively predicting suitable fertilizers based on soil characteristics.

2. Confusion Matrix Analysis

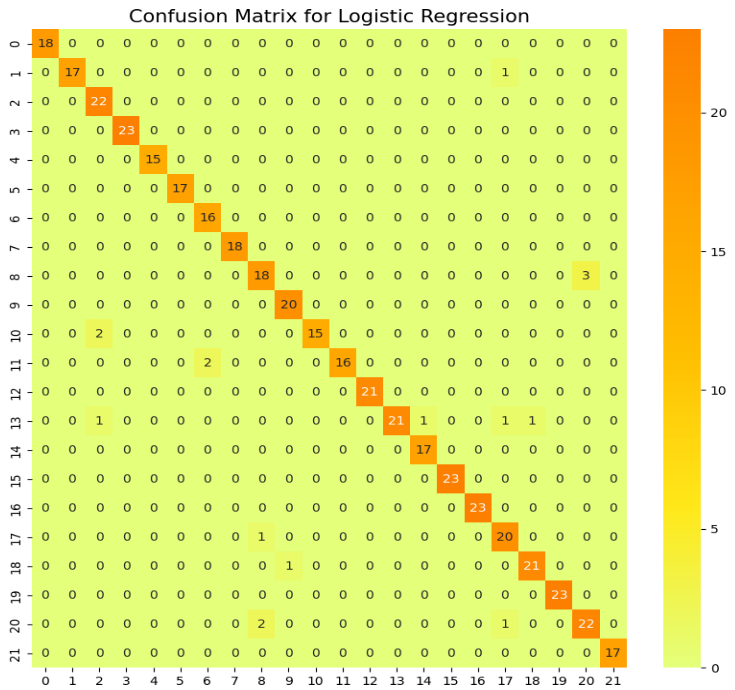
Confusion matrices were used to visualize the performance of both models. They provide insight into the true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) for each class.

Fertilizer Recommendation Model (Random Forest):



- Fertilizer recommendations showed strong accuracy, with very few misclassifications.
- The model performed particularly well in maintaining a high true positive rate (TPR) for both classes.

Crop Recommendation Model (Logistic Regression):



- The majority of predictions are on the diagonal, indicating that most crops were classified correctly.
- Minimal misclassifications (off-diagonal entries) demonstrate the model's high precision and recall.

3. Key Observations

- **High F1 scores** for both models indicate a balance between precision and recall, effectively minimizing false positives and false negatives.
- Logistic Regression’s performance was consistent across classes in crop recommendation.
- Random Forest excelled at fertilizer recommendation by leveraging its ability to handle non-linear relationships and feature importance.

4. Insights from Analysis

- **Crop Model Strengths**: Logistic Regression excelled in capturing linear dependencies in the dataset, offering reliable predictions.

- **Fertilizer Model Strengths**: Random Forest's ensemble learning nature provided robustness to noise and improved generalization.

These results confirm the models' effectiveness and provide a solid foundation for deploying the system in real-world scenarios, supporting farmers in making informed decisions about crop and fertilizer selection.

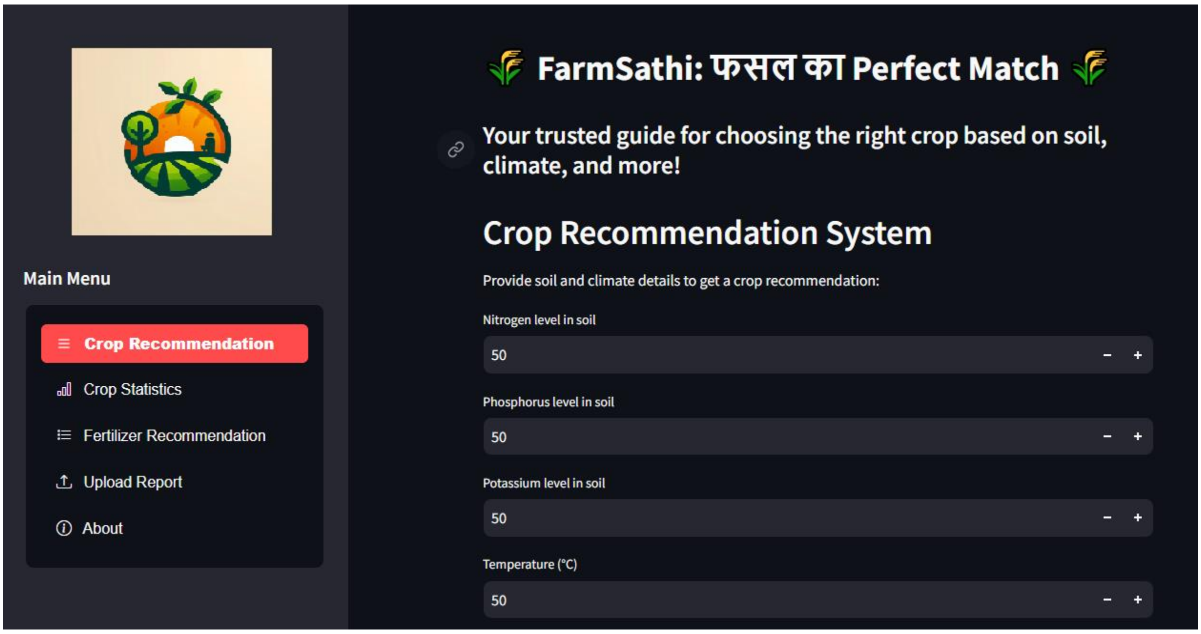
5. Deployment and Usability

The models were deployed in a **Streamlit web application** to ensure ease of use and accessibility for end users.

Key Features:

- **User-Friendly Interface**: Simple input fields for soil properties, climate data, and crop details with easy-to-understand recommendations.
- **Real-Time Predictions**:
  - Crop Recommendation: Suggest the best crop based on environmental and soil data.
  - Fertilizer Recommendation: Recommends fertilizers suited to soil and crop needs.
- **Visualization**: Interactive charts enhance the clarity of recommendations.
- **Scalability**: Designed for future features like pest management and irrigation suggestions.

Application Screenshot:



Future Scope

The project holds significant potential for future enhancements and scalability:

1. **Integration of Additional Data**:
   - Incorporating real-time weather data for dynamic recommendations.
   - Including advanced soil analysis data to improve prediction accuracy.
2. **Expanded Recommendations**:
   - Supporting a broader range of crops and fertilizer types.
   - Adapting the system to include region-specific crops and agricultural practices.
3. **Farmer Assistance Platform**:
   - Developing a comprehensive platform to assist farmers with features like crop disease detection and pesticide recommendations.
   - Providing holistic solutions for farm management and productivity improvement.

These developments aim to make the system more robust, user-centric, and capable of addressing diverse agricultural challenges.

Conclusion

The Crop and Fertilizer Recommendation System demonstrates the potential of AI and machine learning in revolutionizing agriculture. By leveraging Logistic Regression for crop recommendation and Random Forest for fertilizer suggestion, the system achieves high accuracy and reliability, as evidenced by the evaluation metrics. The deployment as a Streamlit web application ensures accessibility and ease of use, enabling farmers and stakeholders to make data-driven decisions efficiently.

The interactive user interface, real-time predictions, and scalability features highlight its practical applicability in addressing agricultural challenges. With planned enhancements like real-time weather data integration, expanded crop recommendations, and advanced farmer assistance features, the system is poised to become a comprehensive solution for modern farming. This project represents a significant step toward sustainable and technology-driven agriculture, paving the way for future advancements in the field.

Reference

- **Scikit-learn Documentation**: Tools and libraries for implementing Logistic Regression and Random Forest models. Available at: https://scikit-learn.org
- **Streamlit Documentation**: Guide for developing and deploying web applications using Python. Available at: https://docs.streamlit.io
- **Kaggle Datasets**: Resources for agricultural datasets used in model training and validation. Available at: https://www.kaggle.com/datasets
- **Google Colab**: Platform used for model training and experimentation with machine learning algorithms. Available at: https://colab.research.google.com
- **Crop Recommendation System using Machine Learning Algorithms;** By Gaurav Chauhan and Alka Chaudhary. Read the paper on IEEE Xplore
- **Crop Recommendations using Machine Learning Techniques;** By Shafiulla Shariff, Shwetha R B, Ramya O G, Pushpa H, and Pooja K R. Read the paper on IJERT