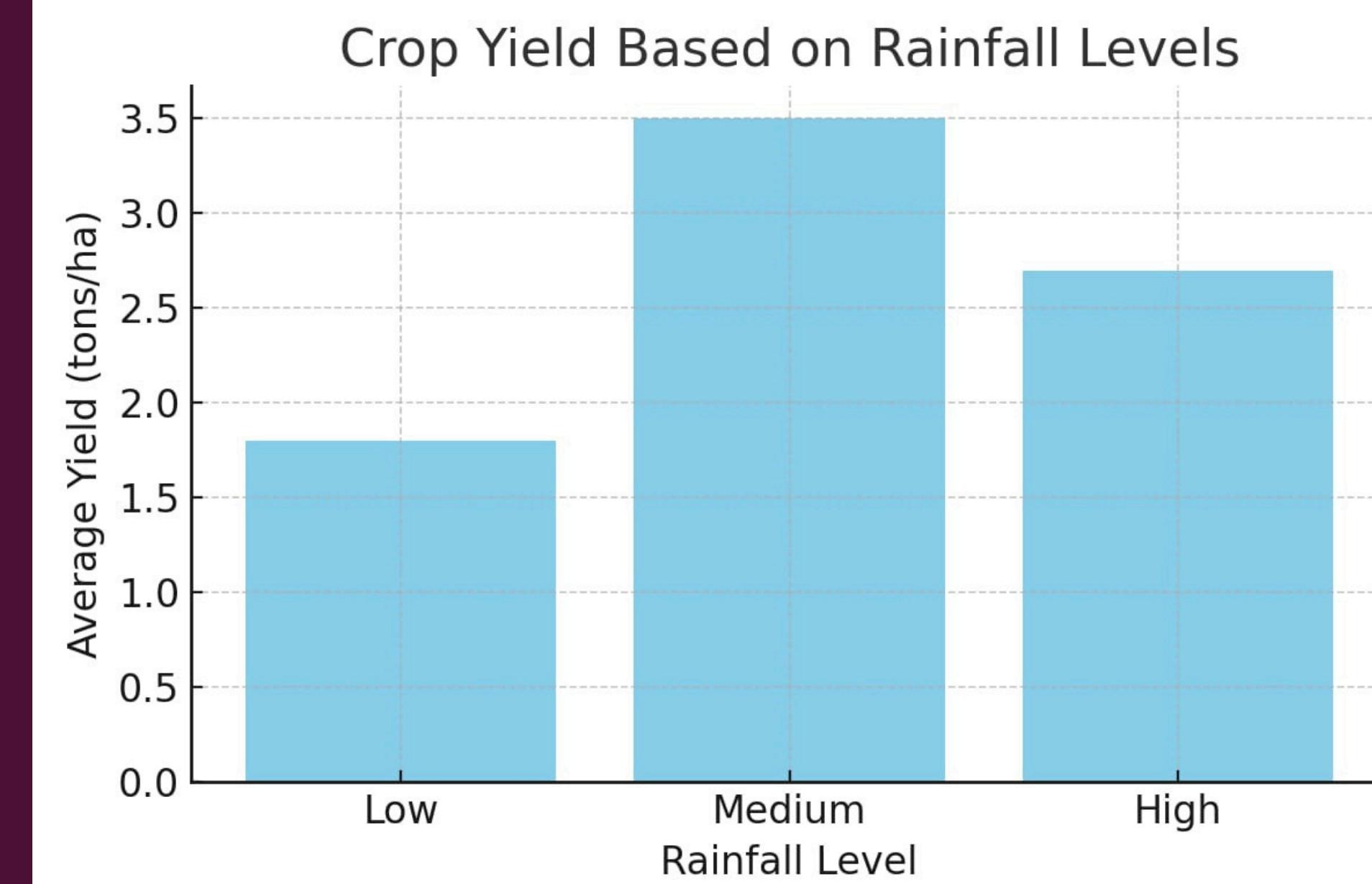


AGRICULTURE CROP YIELD PREDICTION AND PEST DETECTION

By: RA2211003010247 - MONISH
RA2211003010262 - SHREYA
RA2211003010222 - SUSMITHA
RA2211003010201 - SAI NEHA



01. Introduction

In an era defined by data-driven insights, the Agriculture and Crop Prediction Project emerges as a beacon of innovation, aiming to transform traditional farming practices. By leveraging the power of advanced analytics and machine learning, this initiative seeks to provide farmers with the tools they need to optimize crop yields, proactively detect pests, and ensure sustainable agricultural practices.

Technologies Used:

This project was developed using Python, leveraging libraries such as Pandas for data manipulation, Matplotlib and Seaborn for data visualization, and Scikit-learn for implementing machine learning algorithms like Linear Regression and Random Forest. Jupyter Notebook was used as the development environment due to its flexibility in interactive coding and visualization. Additionally, the dataset was sourced from the Indian Government's Agri-Data Portal, and CSV files were used for data storage and preprocessing.

DATASET

Our project utilizes a multi-source dataset derived from the Indian Government's Agri-Data Portal. The primary dataset, `yield_df.csv`, contains comprehensive information on crop yield measured in hectograms per hectare across various countries and crops. It includes key environmental and agricultural variables such as average annual rainfall, pesticide usage in tonnes, and average temperature — all crucial indicators of crop performance. Supporting datasets like `rainfall.csv`, `temp.csv`, and `pesticides.csv` provide region-specific climate and input usage trends. This rich combination of agro-climatic and yield data enables the application of machine learning models for effective crop yield prediction and pest detection analysis.

03. Materials and methods

This project employs publicly available agricultural datasets sourced from the Indian Government's Agri-Data Portal. The primary dataset includes crop yield information alongside critical agro-environmental parameters such as average annual rainfall, pesticide usage (in tonnes), and average temperature. Data preprocessing involved handling missing values, selecting relevant features, and normalizing continuous variables. For the modeling phase, we applied both Linear Regression and Random Forest Regression to predict crop yield. The dataset was split into training and testing sets to evaluate model performance. Random Forest was chosen for its ability to handle non-linear relationships and feature interactions, while Linear Regression provided a simple, interpretable baseline. Model accuracy was assessed using R² scores and visualized through predicted vs. actual yield scatter plots.

02. Objective

Enhance Crop Yield Prediction

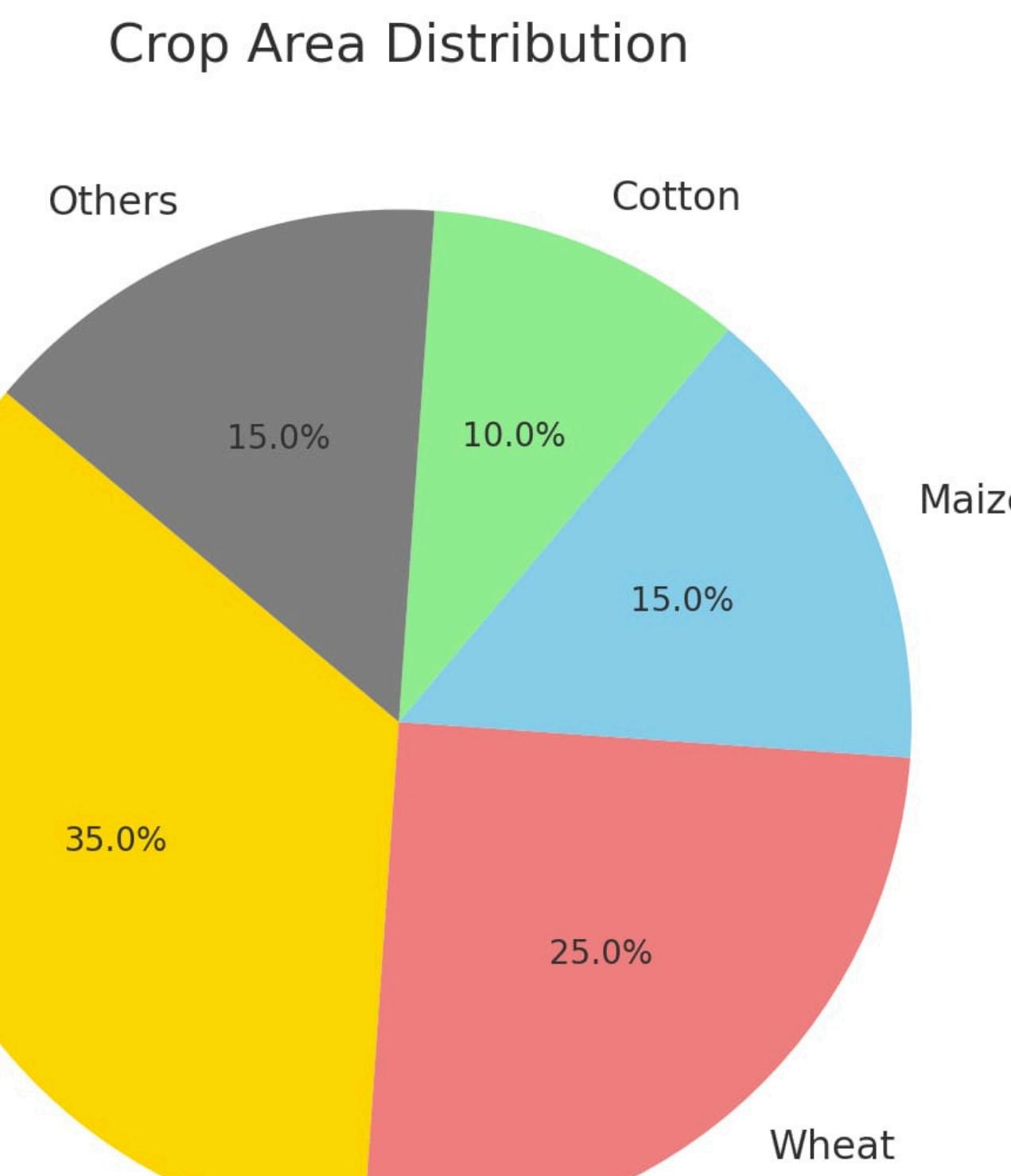
Develop accurate predictive models using agri-data from the Indian government portal to forecast crop yields, enabling farmers to make informed planting decisions.

Enable Early Pest Detection

Implement advanced pest detection systems to identify and mitigate pest infestations, minimizing crop damage and ensuring healthier harvests.

Promote Sustainable Agriculture

Foster environmentally sustainable farming practices through data-driven insights, reducing the reliance on harmful pesticides and promoting efficient resource utilization.



04. Future Research

In the future, this project can be expanded in several directions:

- Pest Detection via Computer Vision: Integrate pest image datasets and use deep learning models like CNNs for real-time pest classification and alerts.
- Remote Sensing & Satellite Data: Incorporate NDVI (vegetation index) and soil moisture data from satellites for enhanced spatial analysis.
- Geospatial Yield Prediction: Develop region-specific models using GIS data for localized yield forecasting.
- Time-Series Forecasting: Use LSTM or ARIMA models for predicting future yields based on temporal data trends.
- Decision Support System: Build an interactive dashboard or app for farmers and policymakers to visualize predictions and receive recommendations.

