FULL STACK DATA SCIENCE SYSTEMS

Crop Planner

Jash Mayur Gosrani - 101423198
Jeet Patel - 101434290
Kaushal Ghelani - 101433516
Satyajeet Singh - 101437086

Introduction

Our goal is to develop a predictive model that recommends the most suitable crops based on soil and environmental conditions.

By analyzing parameters such as nitrogen, phosphorous, potassium levels, temperature, humidity, pH, and rainfall, farmers will receive personalized recommendations for optimal crop selection, ultimately enhancing agricultural productivity and sustainability.

Dataset

We sourced a comprehensive dataset from Kaggle containing information on soil composition, climate conditions, and crop types.

Data fields:

- N - ratio of Nitrogen content in soil
- P - ratio of Phosphorus content in soil
- K - ratio of Potassium content in soil
- temperature - temperature in degree Celsius
- humidity - relative humidity in %
- ph - ph value of the soil
- rainfall - rainfall in mm

- Target Variables ['rice', 'maize', 'chickpea', 'kidneybeans', 'pigeonpeas', 'mothbeans', 'mungbean', 'blackgram', 'lentil', 'pomegranate', 'banana', 'mango', 'grapes', 'watermelon', 'muskmelon', 'apple', 'orange', 'papaya', 'coconut', 'cotton', 'jute', 'coffee']

Model

In our model benchmarking phase, we split the dataset into training and testing sets using a 70-30 ratio. We then experimented with various machine learning algorithms, namely Decision Tree, Gaussian Naive Bayes (NB), and Random Forest. Through evaluation, we observed that Decision Tree achieved an accuracy of 93.18%, Gaussian Naive Bayes achieved 99.39%, and Random Forest achieved 99.24%. Given its simplicity and superior performance, we opted for Gaussian Naive Bayes as our model of choice for the crop recommendation system. This decision was based on its accuracy and effectiveness in handling multi-class classification problems, ensuring reliable and precise crop recommendations for farmers.

Deployment and Evaluation

The trained model was deployed on Heroku using Streamlit, providing an intuitive user interface for farmers to access crop recommendations.

Real-time performance metrics and error logging mechanisms were implemented for live evaluation and monitoring post-deployment.

Future Enhancements

Moving forward, there are several avenues for enhancing the capabilities of our crop recommendation system. Firstly, we aim to incorporate explanations into the model's recommendations, providing farmers with insights into why a specific crop was suggested. By leveraging interpretable machine learning techniques such as decision tree visualization or model-agnostic methods like SHAP (SHapley Additive exPlanations), we can elucidate the underlying factors driving each recommendation, such as soil composition, climate conditions, and historical crop yields. Additionally, we plan to expand the functionality of the model to offer multiple crop suggestions for a given set of input parameters. This enhancement would empower farmers with a diverse range of options, enabling them to make more informed decisions based on their preferences, market demands, and agronomic considerations. By continually refining and expanding the capabilities of our model, we strive to provide farmers with comprehensive and actionable insights to optimize their crop selection strategies and enhance agricultural productivity.

Appendix

- Code repository
- Live demo