Crop Planner

Business Case

Background

In today's agriculture landscape, farmers face numerous challenges including unpredictable weather patterns and soil variability. Data-driven solutions are crucial for addressing these challenges.

Our project builds upon existing datasets on rainfall, climate, and fertilizer data, augmenting them to create a robust foundation for predictive modeling.

Project Definition

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.

Project Implementation

Our implementation plan follows a structured timeline, with milestones and responsibilities clearly defined.

1. Project Kickoff	6. Documentation
2. Data Collection and Preparation	7. Deployment
3. Model Development	8. Training and Support
4. User Interface Design	9. Evaluation and Iteration
5. Integration and Testing	10. Conclusion and Handover

Business Requirements

Accuracy, usability, and scalability are paramount, ensuring the model is both effective and accessible to farmers.

Compliance with data regulations ensures ethical data management practices, building trust and reliability within the agricultural community.

Model and Deployment

After careful consideration, we've chosen Gaussian Naive Bayes as our preferred option. This algorithm is suitable for this task as it assumes that the features follow a Gaussian (normal) distribution, which is often a reasonable assumption for many real-world datasets. It's a simple and efficient algorithm which aligns with our project goals. Utilizing Streamlit, we'll develop a user-friendly interface and deploy it on Heroku for farmers to access crop planner seamlessly.

Evaluation

We've evaluated multiple approaches for model development, considering factors such as algorithm complexity, tool availability, and deployment feasibility.

Options range from machine learning algorithms to traditional statistical methods, each with its pros and cons. Our decision prioritizes accuracy, interpretability, and scalability

Benefits and Risks

The benefits of our crop planner project are manifold. Farmers gain access to tailored recommendations, improving crop yields and profitability. Resource efficiency reduces input costs and environmental impact. Ultimately, the project enhances food security and rural livelihoods.

While the project offers significant benefits, risks must be acknowledged. Data quality issues, model accuracy, and adoption barriers pose challenges. Mitigation strategies include rigorous testing, stakeholder engagement, and ongoing monitoring to ensure project success.

Financial Analysis

- Cost Estimation: Evaluate project expenses including data collection, model development, and software infrastructure.
- Revenue Projection: Estimate revenue gains from increased crop yields facilitated by the crop planner system.
- Return on Investment (ROI) Calculation: Determine the project's ROI by comparing projected revenue with initial investment.
- Funding Options: Explore grants, partnerships, and subsidies to secure financial resources.
- Financial Risk Assessment: Identify and mitigate risks such as cost overruns and revenue uncertainty.

Resources Required

- Personnel: Assemble a team of data scientists, developers, and agricultural experts.
- Technology Infrastructure: Ensure access to cloud resources and development tools.
- Data Resources: Utilize comprehensive datasets for model training and validation.
- Training and Capacity Building: Develop programs to educate users on the crop planner system.
- Stakeholder Engagement: Foster collaboration with farmers, organizations, and government agencies.

Methodology

Data Collection and Preprocessing

To initiate the process, from Kaggle we sourced a comprehensive dataset containing information on soil composition, climate conditions, and crop types. This dataset serves as the foundation for training our predictive model.

Upon loading the dataset, we embarked on a crucial step of data preprocessing to ensure its quality and compatibility with our model.

This involved converting non-numeric columns to numeric where possible and handling any missing values to create a clean and standardized dataset.

About Dataset

This dataset was build by augmenting datasets of rainfall, climate and fertilizer data.

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

Model Selection and Training

After exploring various machine learning algorithms including K-Nearest Neighbors (KNN), Random Forest, Decision Tree, etc but we selected Gaussian Naive Bayes as the model of choice for its simplicity and effectiveness in handling multi-class classification problems, such as our crop planner.

Leveraging the preprocessed dataset, we split the data into training (70%) and testing (30%) sets to facilitate model training while ensuring an unbiased evaluation of its performance.

Model Evaluation and Deployment

With the model trained, we proceeded to evaluate its effectiveness in predicting crop recommendations based on input parameters. We employed classification metrics such as precision, recall, and F1-score to assess the model's performance across multiple crop classes.

We transitioned to the deployment phase, leveraging Streamlit to create an intuitive and user-friendly interface. This interface allows farmers to input relevant soil and environmental parameters and receive real-time crop recommendations generated by the trained model.

DEMO


Thank You