Optimizing Fertilizer Usage for Sustainable Farming

Chua Jing Jie
12th June 2024

Problem Statement

- Challenging for farmers to determine the optimal amount of fertilizer to use for their crops.
  - Too much fertilizer = soil degradation and wasted resources
  - Too little fertilizer = poor crop growth and low yield
- Building a decision support system to provide optimal fertilizer amount to use based on:
  - Soil conditions
  - Weather conditions
  - Crop types
- Farmers can input these easily obtainable data into the system to receive optimal fertilizer amount to use to:
  - Increase crop yield
  - Reduce costs
  - Minimise environmental impact

Data Schema for Dummy Data

Feature	Type	Description
soil_color	String	Qualitative measure of soil color (e.g., dark brown, reddish)
soil_ph	Float	Soil pH level
soil_n	Float	Nitrogen content in soil (ppm)
soil_p	Float	Phosphorus content in soil (ppm)
temp	Float	Current temperature (°C)
rainfall	Float	Recent rainfall (mm)
forecast_temp	Float	Forecasted temperature (°C)
forecast_rainfall	Float	Forecasted rainfall (mm)
crop_type	String	Type of crop being grown
plant_health	String	Qualitative measure of plant health (e.g., healthy, yellowing, wilting)
optimal_fertilizer_amount	Float	Target value: Optimal fertilizer amount (kg/ha)

Assumptions:

- 10 Features, 1000 Rows of dummy data generated
- Features like soil content, weather, crop type, and plant health affect optimal fertilizer amount
- Farmers can collect these data easily to get recommendations
- Low percentage of outliers and missing data to mimic real-world imperfections
- Temperature and rainfall values are based on Indonesia's tropical climate
- Crop types are limited to common crops grown in Indonesia (wheat, corn, rice)
- Plant health conditions are randomly assigned as healthy, yellowing, or wilting.
- Fertilizer amount calculated with a simplified formula based on input features

Model Development Process

Simplified model development process to predict optimal fertilizer usage

Process:

1. **Data Splitting**: dummy data split into 67% train and 33% test sets
2. **Preprocessing Pipeline**
    a. Fill missing values with median
    b. Scale numeric features
    c. Encoded categorical features
2. **Model training:** Trained an XGBRegressor model using preprocessed train data
3. **Evaluation**: RMSE of 4.28 when predicting on test set

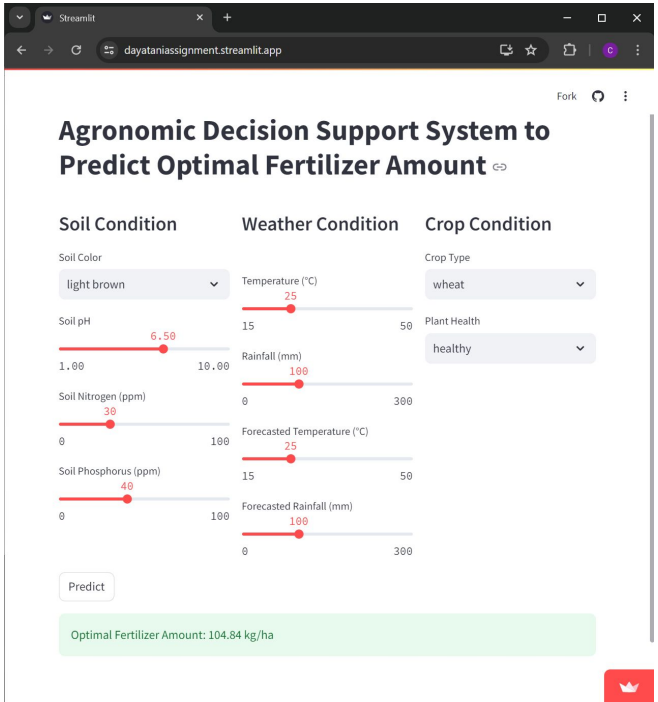
Recommendations for Farmers (2 ways to give practical recommendations)

1. Give **general recommendations** for optimal fertilizer amounts by using the model to predict based on pre-defined scenarios:
   a. Yellowing wheat with low rainfall and low soil nutrient levels: Use 132.80 kg/ha of fertilizer
   b. Healthy rice with high rainfall and high soil nutrient levels: Use 105.19 kg/ha of fertilizer
   c. Wilting wheat with very low nutrient levels and low rainfall: Use 136.64 kg/ha of fertilizer

---

2. Give **tailored recommendations** to farmers by predicting based on the specific data that the farmer provided.

- Created a simple Streamlit web app to host the model to demonstrate how farmers can access the model via web/mobile apps:
  - Link to web app: https://dayataniassignment.streamlit.app/
  - Farmers enter soil, weather, crop data
  - Receive predicted optimal fertilizer amount



Thank You