Optimized Fertilizer Quantity Recommendation using Indian Soil Dataset via Machine Learning

Piyush Gupta
IIITDM Kurnool
gguptapiyush45@gmail.com

Abstract

Precision in fertilizer application is crucial for sustainable agriculture in India. This work presents a machine learning-based system for recommending optimal quantities of Urea, DAP, and MOP using geospatial soil characteristics derived from raster datasets provided by ISRO's NICES portal. The raster data is converted to structured form and processed with an XGBoost classifier. The model demonstrates 100% classification accuracy and is integrated into a Gradio web interface for real-time recommendation.

1 Problem Statement & Motivation

Indian farmers often lack access to soil-specific fertilizer guidance, leading to under- or overuse. Traditional recommendations ignore soil texture, carbon density, and depth. This project aims to create a machine learning pipeline that personalizes fertilizer recommendations based on geospatial soil maps, enabling sustainability and better yield optimization.

2 Dataset Source & Composition

Source: ISRO Bhuvan – NICES Soil Repository
Original Format: Raster maps (.asc, .hdr, .xml, .jpg)
Processed Format: Structured tabular data (via rasterio)
Features Extracted:

- Soil Texture: Sandy, Loamy, Clayey, Clayskeletal
- Organic Carbon Density: Mean values per coordinate
- Soil Depths: Aggregated across 6 bands (0–25 to 150–200 cm)

Final dataset: Over 430,000 geotagged data points with soil features and derived labels.

3 Label Generation Strategy

Due to absence of labeled fertilizer usage, we implemented a rule-based logic derived from agronomic insights:

- Organic carbon thresholds define Urea recommendation levels.
- DAP and MOP recommendations use average soil depth and adjusted thresholds.
- Soil texture helps refine the fertilizer demand class.

Labels for each fertilizer: *Low, Medium, High*

4 Feature Engineering

Input Features (X):

- Encoded SoilTexture
- OrganicCarbonDensity (float)
- Average Depth (mean of 6 layers)

Target Variables (y): Urea_Rec, DAP_Rec, MOP_Rec

5 Model Selection & Evaluation

We used **XGBoost Classifier** for its performance and feature importance insights. The dataset was split 80:20 for training and testing.

Metrics Achieved:

- Accuracy: 100% for all three fertilizers
- Precision, Recall, F1-score: All perfect

6 Deployment

The final model was deployed in Google Colab using the Gradio interface.

Inputs: Soil Texture, Organic Carbon, Soil Depth
Outputs: Urea, DAP, and MOP recommendations in Low/Medium/High

7 Conclusion & Future Work

We developed a robust ML pipeline for fertilizer recommendation using real geospatial data. It is scalable, interpretable, and suitable for integration with real-time field systems.

Future Directions:

- Integration with crop type and real-time weather
- Interactive GIS maps for farmers
- Learning from crop yield feedback to refine predictions

Author

Piyush Gupta
Machine Learning Student, IIITDM Kurnool
gguptapiyush45@gmail.com

Supervisor

Dr. Vishesh P. Gaikwad
Assistant Professor, Department of CSE
Sardar Vallabhbhai National Institute of Technology (SVNIT), Surat