AgriAssist

Introduction:

Empowering farmers to fine-tune the nutrient content in their soil, resulting in healthier crops and the possibility of increased yields. Additionally, it will aid in promoting sustainable farming methods by minimizing the use of excess fertilizers, which in-turn helps to cut down on farmers expenses.

Problem Statement:

Agriculture serves as one of the pillars of our country. Commonly, when crops need chemical treatment, farmers typically consult nearby vendors for purchasing these chemicals. Unfortunately, owing to certain unethical business practices, these vendors frequently provide an excess of chemicals, surpassing the requirements. This not only raises the expenses associated with spraying but also alters the uniformity of the soil's nutrient composition. To address this issue, our project aims to leverage soil data collected via sensors. We plan to process this data through machine learning (ML) models to accurately identify nutrients that are lacking in the soil and subsequently advise farmers on the specific nutrients needed. At this stage, our focus will be primarily on developing and fine-tuning the ML models. In the future, we intend to enhance our application to directly gather data from the sensors. This approach is designed to equip farmers with precise, data-driven insights, reducing their dependence on external vendors and promoting more cost-effective and sustainable farming practices.

Data used and description:

We have used Soil nutrients data as our dataset, where each row in the data represents a sample of soil and the nutrient value observed while sampling. There are 12 indicators of soil nutrients we have used in our dataset; a tabular description of each nutrient is given below:

- Macronutrients, such as N, P, K, and S, are measured in units of kg/ha.
- Micronutrients, including Zn, Fe, Cu, and Mn, are quantified in mg/kg.
- Organic Carbon (OC) levels are expressed in g/kg.
- Electrical Conductivity (EC) values are denoted in mS/m.

Column	Description
N	Nitrogen

P	Phosphorus
K	Potassium
pH	indication of the acidity or alkalinity of the soil
EC	Soil electrical conductivity (EC) measures the amount of soluble salts in soil.
OC	Organic Carbon
S	Sulphur
Zn	Zinc
Fe	Iron
Cu	Copper
Mn	Manganese
B	Boron

Project Artifacts:

We have incorporated three Python files as follows:

- **Group-16_AgriAssist_EDA.ipynb**: This notebook includes exploratory data analysis (EDA) and the distributions for each feature in the dataset.
- **Group-16_AgriAssist_Model.ipynb**: This notebook encompasses PCA, RandomForest, and other techniques, alongside Isomap and t-SNE for data visualization.
- **app.py**: This script integrates PCA, Random Forest, and Streamlit deployment code.
- **Group-16_app.ipynb:** This document is a conversion of app.py into a .ipynb file format for submission purposes, as the submission guidelines do not permit .py files. When deploying on Streamlit, ensure the file is named app.py

Dataset -> Uploaded to GitHub repository

GitHub repository for AgriAssist:

https://github.com/Dhanush-Garrepalli/AgriAssist

Streamlit URL for AgriAssist:

https://agriassist.streamlit.app/

Main findings & outcomes

From our investigation, it's clear that there is both an overapplication and underutilization of pesticides. It's crucial for farmers to be aware of the specific mineral surpluses or deficits in their soil to enhance fertility.

Regarding our dataset, it has been compiled from various research publications and soil data websites, encompassing different years and climatic conditions. For our analytical purposes, we focused on 12 essential nutrients, setting aside weather-related data for these instances. Concerning our model, we employed PCA and t-SNE techniques to pinpoint the six most significant nutrients, which we then utilized for model-based predictions and soil fertility analyses based on user inputs.

Our model is designed to assess soil fertility based on user-provided nutrient levels, specifically P, pH, EC, Mn, N, and Cu. Through a user interface, when values for these nutrients are entered, the model conducts an in-depth analysis of each nutrient's level and offers a comprehensive prediction of the soil's overall fertility.

Scope for improvement

During the development of our model, we discovered that achieving optimal crop yield in the agriculture sector hinges on two primary factors: weather conditions and soil fertility. Presently, our model operates under stable weather conditions. Incorporating weather conditions and suitable crops presents an opportunity for enhancement. However, crafting a universal template for agriculture remains a challenging endeavor due to the variability of these factors.

In real-time scenarios, soil nutrient levels are determined using soil test kits or electronic soil meters.

Conclusion

Utilizing nutrient level inputs supplied by users through the UI, we developed a model capable of predicting overall soil fertility. It assesses each nutrient individually, providing recommendations on surplus or deficiency for specific nutrients.

Project Contributions:

Problem Statement: Defined by Nehal and Dhanush.

Exploratory Data Analysis (EDA): Conducted by Shashank and Karan.

Data Visualization: Managed by Shashank and Nehal.

Principal Component Analysis (PCA): Executed by Karan and Nehal.

Anomaly Detection: Implemented by Dhanush and Anjana.

Streamlit Deployment: Completed by Anjana and Dhanush.

Documentation/Reporting: Prepared by Shashank, Anjana, and Karan.

References:

https://www.researchgate.net/figure/Soil-fertility-ratings-for-available-nutrients_tbl2_315767883

https://www.isric.org/explore/soilgrids

MLUL2 Group 16 Assignment by:

Karan Singh (12310034)

Sugunaanjana Vasireddy (12310086)

Garrepalli Dhanush (12310091)

Nehal Singh (12310102)

Shashank Thakur (12310109)