

Project Title: Agricultural Plant Classification Using Machine Learning

Problem Statement:

This project focuses on developing a machine learning model to classify agricultural plants based on various dimensional and shape factors. The primary objective is to build an accurate and efficient classification system by exploring different data preprocessing techniques, visualization methods, and classification algorithms.

Project Goals:

1. **Data Preprocessing:**
 - Clean and prepare the dataset by handling missing values, scaling features, and encoding categorical variables.
 - Ensure the data is optimized for effective analysis and model building.
2. **Data Visualization:**
 - Create insightful visualizations to explore the dataset, understand feature distributions, and identify relationships between variables.
 - Detect patterns, trends, and potential correlations to inform model selection.
3. **Model Development:**
 - Apply multiple classification algorithms such as Decision Trees, Random Forests, Support Vector Machines (SVM), and K-Nearest Neighbors (KNN) to build predictive models.
 - Optimize hyperparameters to enhance model performance.
4. **Model Evaluation:**
 - Evaluate the models using performance metrics like Accuracy, Precision, Recall, F1 Score, and ROC AUC.
 - Identify the most effective algorithm for classifying the agricultural plants.
5. **Comparison & Insights:**
 - Compare the results of different classification models to determine the most suitable approach.
 - Provide insights and recommendations based on the evaluation outcomes.

Tools & Technologies Used:

- Python
- Jupyter Notebook
- Pandas, NumPy (for data preprocessing)
- Matplotlib, Seaborn (for data visualization)
- Scikit-learn (for machine learning models and evaluation)

This project highlights my approach to solving real-world classification problems, focusing on data preprocessing, model selection, and performance evaluation to deliver actionable insights.