Project Title:

Comparative Study of Image Classification Using Decision Tree, Naive Bayes, and Feedforward Neural Networks

Project Overview:

This project aims to implement and compare different machine learning models for the task of image classification. Student or (group of 2) will work with a dataset of at least 500 images, organized into three or more classes (e.g., religious symbols, traffic signs, leaf types, animals types). Each image should be resized to a fixed dimension (e.g., 32×32 or 64×64) and flattened into a 1D feature vector for model input.

The models to be implemented are:
1. Naive Bayes Classifier – fast and effective for high-dimensional data like pixel values.
2. Decision Tree Classifier – interpretable and suitable for rule-based visual classification.
3. Feedforward Neural Network (MLPClassifier) – optional model to explore deeper representations.

Students will train and evaluate these models and analyze their accuracy, efficiency, and ability to generalize to unseen data.

Project Objectives:

- - Prepare and preprocess an image dataset for classification.
- - Implement Naive Bayes and Decision Tree models.
- - Optionally implement a Multi-Layer Perceptron (MLP) using MLPClassifier.
- - Evaluate models using metrics such as accuracy, precision, recall, F1-score, and confusion matrix.
- - Analyze and compare the strengths and limitations of each model.

Dataset Requirements:

• Minimum of 500 labeled images, divided into at least three categories.
• Images must be:
  - Resized (e.g., 32×32 or 64×64).
  - Flattened into 1D vectors for model input.
• Example dataset themes: religious icons, traffic symbols, types of leaves.

Models to Implement:

1. Naive Bayes Classifier:
   Use simple pixel-level features or statistical summaries (mean, variance). Evaluate as a baseline model.
2. Decision Tree Classifier:
   Build a decision tree that classifies images using pixel intensity values. Visualize and analyze decision paths.

3. Feedforward Neural Network (optional):
   Use Scikit-learn's MLPClassifier with 1–2 hidden layers. Compare performance with the simpler models.

Project Deliverables:

- - Source Code for all models with clear documentation.
- - Final Report (8 pages), including:
  - Introduction and dataset description.
  - Detailed explanation of each model.
  - Evaluation results with accuracy metrics and confusion matrices.
  - Comparative analysis and discussion of which model performs best and why.