Lab Assignment 6: Classification Decision Tree using the C4.5 Algorithm

Duration: 2 Hours Date: 02-04-2025

Dataset: WineQT

Objective:

Implement the C4.5 algorithm to classify wine quality as "Good" or "Not Good" using a decision tree based on the dataset's features.

Instructions:

- **Data Preprocessing:**
  - Load the Wine Quality dataset.
  - Convert quality scores into binary labels (Good ≥ 7, Not Good < 7).
  - Split data into training (80%) and testing (20%) sets.
- **CART Algorithm Implementation:**
  - Use **scikit-learn** to implement CART (criterion: **Gini Index** or **Entropy**).
  - Build a decision tree, limiting depth to avoid overfitting.
- **Model Training and Evaluation:**
  - Train the model on the training set.
  - Evaluate performance using metrics: **Accuracy**, **Precision**, **Recall**, **F1-Score**, and **Confusion Matrix**.
  - Visualize the decision tree.
- **Conclusion:**
  - Analyze and summarize results based on evaluation metrics and tree structure.
  - Summarize your findings, discussing the strengths and weaknesses of your model.

Report: 1-2 page summary with visualizations and model performance.