Lab Exercises 19

Date: 11 Nov 2025

Implementation of Decision Tree Classifier

from sklearn.tree import DecisionTreeClassifier, plot_tree

19. Implement a Decision Tree Classifier to predict the survival of passengers using the Titanic dataset.
    a. Perform necessary preprocessing (handle missing values, encode categorical data).
    b. Train and visualize a Decision Tree Classifier to predict the Survived column using plot_tree() or graphviz.
    c. Evaluate the model using appropriate classification metrics.

Performance Evaluation: Answer the following five evaluation questions based on your model results:

1. **What is the overall accuracy of your Decision Tree Classifier on the test set?**

   (Use accuracy_score from sklearn.metrics)

2. **Construct and display the confusion matrix. How many passengers were correctly and incorrectly classified as survivors/non-survivors?**

   (Use confusion_matrix and ConfusionMatrixDisplay)

3. **Compute and interpret the Precision, Recall, and F1-Score of the classifier.**

   (Use classification_report)

4. **Display the feature importance scores. Which features most influence survival prediction?**

   (Use model.feature_importances_)

5. **Visualize the Decision Tree and explain one of the decision paths (e.g., how a "female passenger in 1st class" is classified).**

   (Use plot_tree() and analyze one branch of the tree)