Machine Learning Assignment: Decision Trees

# Step 1: Data Understanding

1. 1. Load the dataset into Python using pandas.
2. 2. Display the first 5 rows and describe each feature.
3. 3. Check for missing values and handle them if any.

# Step 2: Data Preprocessing

1. 4. Convert categorical variables (Extracurricular, Internet\_Access) into numerical form.
2. 5. Split the dataset into features (X) and target (y).
3. 6. Perform train-test split (70% train, 30% test).

# Step 3: Model Building

1. 7. Train a Decision Tree Classifier on the training set.
2. 8. Print the learned decision rules (tree structure).

# Step 4: Model Evaluation

1. 9. Predict on the test set.
2. 10. Evaluate performance using accuracy, confusion matrix, and classification report.

# Step 5: Model Tuning

1. 11. Experiment with max\_depth and criterion (gini vs entropy). Compare results.
2. 12. Plot the decision tree and interpret which features are most important.

# Step 6: Insights & Conclusion

1. 13. Which features influence student performance the most?
2. 14. If you were a teacher, how would you use these insights to help students improve?