Classification Report - Decision Tree Classifier
(Questions & Answers)

- Q1. What is the precision for class 0, and what does it indicate?

A1. The precision for class 0 is 0.90. This means that out of all the samples predicted as class 0, 90% were actually class 0. It indicates a relatively low false positive rate for this class.

- Q2. What is the recall for class 1, and what does it signify?

A2. The recall for class 1 is 0.84. This means that 84% of the actual class 1 instances were correctly identified by the Decision Tree model. It shows the model's effectiveness in detecting true positives for class 1.

- Q3. What is the F1-score for class 1 and why is it important?

A3. The F1-score for class 1 is 0.83. It combines precision and recall into a single metric, giving a balance between them. It's especially useful in cases of class imbalance.

- Q4. What does the "support" value indicate for each class?

A4. The support indicates the actual number of samples for each class in the dataset: 85 for class 0 and 49 for class 1. It's used to weigh the overall performance metrics.

- Q5. What is the accuracy of the Decision Tree model?

A5. The model has an accuracy of 0.87, meaning it correctly predicted 87% of the total 134 samples.

- Q6. How are macro average and weighted average different here?

A6. Macro average is the simple average across classes (precision: 0.86, recall: 0.87, F1: 0.86) without considering class size. Weighted average considers the number of instances per class, giving a more realistic average (all metrics: 0.87) for imbalanced datasets.

- Q7. Why is weighted average preferred in real-world imbalanced datasets?

A7. Since class 0 has more instances (85) than class 1 (49), the weighted average reflects model performance more accurately by accounting for the class distribution.

Confusion Matrix :

Confusion Matrix Diagram:

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Confusion matrix creates a N X N matrix, where N is the number of classes or categories that are to be predicted. Here we have N = 2, so we get a 2 X 2 matrix. Suppose there is a problem with our practice which is a binary classification. Samples of that classification belong to either Yes or No. So, we build our classifier which will predict the class for the new input sample. After that, we tested our model with 165 samples, and we get the following result.

Formulas for Classification Metrics

- Precision:

Precision = TP / (TP + FP)

- Recall:

Recall = TP / (TP + FN)

- F1-score:

F1 = 2 * (Precision * Recall) / (Precision + Recall)

- Accuracy:

Accuracy = (TP + TN) / (TP + TN + FP + FN)

- Macro Avg:

Macro Avg = Average(metric over all classes)

- Weighted Avg:

Weighted Avg = Σ(metric_i * support_i) / Σ(support_i)