Classification Report - Questions & Answers

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

A1. The precision for class 0 is 0.93. This means that out of all the samples predicted as class 0, 93% were actually class 0. It indicates a 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 0.88. This means that 88% of the actual class 1 instances were correctly identified by the model. It shows how well the model is detecting class 1 cases (true positives).

- Q3. What is the F1-score for class 0 and why is it important?

A3. The F1-score for class 0 is 0.92. It is the harmonic mean of precision and recall, balancing both false positives and false negatives. A high F1-score indicates a strong balance between precision and recall.

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

A4. The support value shows how many actual instances of each class exist in the dataset: 85 for class 0 and 49 for class 1. It helps understand class distribution and is important for interpreting averages.

- Q5. What is the overall accuracy of the model, and how is it computed?

A5. The overall accuracy is 0.90, meaning 90% of all predictions (correct predictions / total samples) were accurate.

- Q6. What is the difference between macro and weighted averages in this report?

A6. Macro avg is the unweighted average of the metrics for both classes. It treats all classes equally:
- Precision: 0.89
- Recall: 0.90
- F1-score: 0.90

Weighted avg considers class imbalance by averaging the metrics weighted by support:
- Precision: 0.90
- Recall: 0.90
- F1-score: 0.90

- Q7. Why might weighted average be more reliable than macro average in this case?

A7. Since class 0 has more samples (85) than class 1 (49), the weighted average accounts for this imbalance, giving a more representative performance measure of the model on the entire dataset.

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.

There are 4 terms you should keep in mind:

1.**True Positives:** It is the case where we predicted Yes and the real output was also Yes.

2.**True Negatives:** It is the case where we predicted No and the real output was also No.

3.**False Positives:** It is the case where we predicted Yes but it was actually No.

4.**False Negatives:** It is the case where we predicted No but it was actually Yes.

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)