PRESENTATION

By Isabel Mercado

1. CONFUSION MATRIX

A Confusion Matrix is a table used to evaluate the performance of a classification model.

# 1. CONFUSION MATRIX

- It summarizes the predictions against the actual outcomes.
- It creates an N X N matrix, where N is the number of classes or categories that are to be predicted.
- Suppose there is a problem, which is a binary classification, then N=2 (Yes/No).
- It will create a 2x2 matrix.

1. CONFUSION MATRIX

- True Positives: It is the case where the model predicted Yes and the real output was also yes.
- True Negatives: It is the case where the model predicted No and the real output was also No.
- False Positives: It is the case where the model predicted Yes but it was actually No.
- False Negatives: It is the case where the model predicted No but it was actually Yes.

1. CONFUSION MATRIX
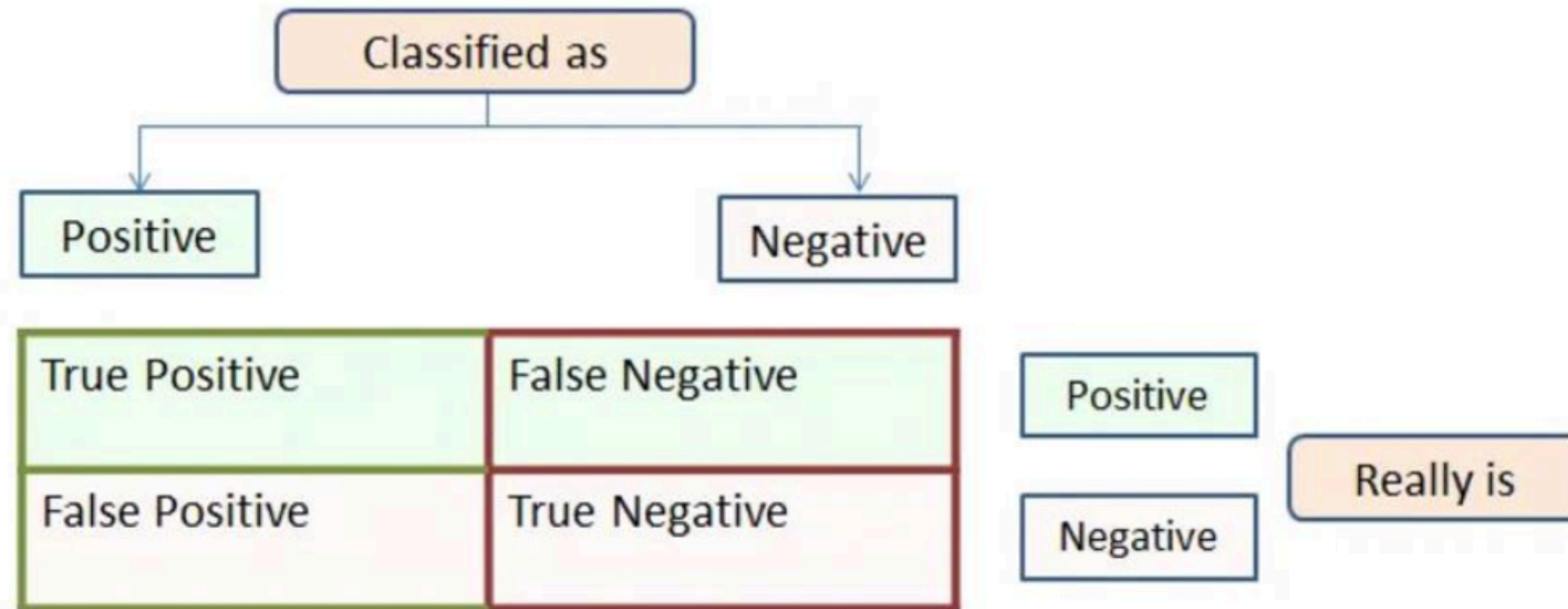
2. PRECISION AND RECALL

Precision measures
"What proportion of predicted Positives is truly Positive?"
Precision = (TP)/(TP+FP).
Precision should be as high as possible.
Recall measures
"What proportion of actual Positives is correctly classified?"
Recall = (TP)/(TP+FN)

3. F1-SCORE

- *A good F1 score means that you have low false positives and low false negatives*, so you're correctly identifying real threats, and you are not disturbed by false alarms.

- An F1 score is considered perfect when it is 1, while the model is a total failure when it is 0.

- **F1 = 2* (precision * recall)/(precision + recall)**

4. ACCURACY

Accuracy = Number of correct predictions / Total number of predictions

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

2.4.2. Given a confusion matrix, calculate Precision, Recall, F1 score and Accuracy

Confusion Matrix:

$$\begin{bmatrix} TN = 50 & FP = 10 \\ FN = 5 & TP = 35 \end{bmatrix}$$

From the confusion matrix:

True Positives (TP) = 35, True Negatives (TN) = 50, False Positives (FP) = 10, False Negatives (FN) = 5

1. Precision

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{35}{35 + 10} = \frac{35}{45} \approx 0.778 \, (77.8\%)$$

Precision indicates how many of the predicted positives are correct.

2. Recall (Sensitivity or True Positive Rate)

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{35}{35 + 5} = \frac{35}{40} = 0.875 \ (87.5\%)$$

Recall shows how many of the actual positives were correctly predicted.

3. F1 Score

$$F1 \text{ Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Substituting values:

$$F1 \text{ Score} = 2 \times \frac{0.778 \times 0.875}{0.778 + 0.875} = 2 \times \frac{0.681}{1.653} \approx 0.823 \, (82.3\%)$$

The F1 Score balances Precision and Recall.

4. Accuracy

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Substituting values:

$$\text{Accuracy} = \frac{35 + 50}{35 + 50 + 10 + 5} = \frac{85}{100} = 0.85 \, (85\%)$$

Accuracy indicates the overall correctness of predictions.

Summary of Metrics:

Precision: 77.8%

Recall: 87.5%

F1 Score: 82.3%

Accuracy: 85%

THANK YOU