# Classification Metrics

## Accuracy [1]

Accuracy measures the proportion of correctly classified instances versus the total instances:

$$accuracy = \frac{corrects}{total\ instances}$$

Where 'corrects' is the total number of true positives (TP) and negatives (TN), whereas 'total instances' is the total number of set's instances, that is, TP, TN as well as false positive (FP) and negative (FN).

This is a well-intuited classification metric that can be quickly calculated. However, it can be misleading in cases where the classes are highly imbalanced.

## Balanced Accuracy

Balanced accuracy is a metric designed to handle imbalanced datasets, where the number of instances in each class is not equal. Unlike standard accuracy, which can be misleading when one class dominates, balanced accuracy provides a more reliable measure by taking both classes into account. Here is the mathematical definition:

$$balanced\ accuracy = \frac{1}{2}\left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right)$$

This formula computes the average recall for both the positive and negative classes, making it less sensitive to class imbalance.

## Recall or Sensitivity (True Positive Rate) [2]

Recall is a metric that measures how well the model identifies positive instances in the dataset. It calculates the proportion of actual positives that are correctly identified by the model. It is particularly useful when the goal is to identify as many positive instances as possible, even at the cost of some false positives. For example, in medical diagnostics (such as cancer detection), it is more critical to identify all possible positive cases, even if it means incorrectly classifying some healthy individuals as sick. Here is the mathematical definition:

$$recall = \frac{TP}{TP + FN}$$

Tip: It answers in the question: “How many of the actual positives did you predict?”

# Precision [2]

Precision is a metric that measures the accuracy of the positive predictions made by the model. It calculates the proportion of predicted positive instances that are actually positive. In other words, it focuses on how many of the positive predictions made by the model were correct. Here is the mathematical definition:

$$precision = \frac{TP}{TP + FP}$$

The formula calculates the fraction of predicted positive instances that are correct. A precision of 1 (or 100%) means that every instance predicted as positive was indeed positive, while a precision of 0 means that none of the predicted positives were actually positive.

Tip: It answers in the question: “How many of the predicted positives are actual positives?”

# F1 Score [3]

The F1 Score is the harmonic mean of Precision and Recall. It combines both metrics into a single value, balancing the trade-off between them. This is particularly useful when you need a single score that captures both false positives and false negatives. Here is the mathematical definition:

$$F1\ score = \frac{2 * precision * recall}{precision + recall}$$

# Matthews Correlation Coefficient (MCC) [4]

The Matthews Correlation Coefficient (MCC) is a metric used for binary classification that provides a measure of the quality of binary classifications. It takes into account all four quadrants of the confusion matrix: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). MCC is especially useful for imbalanced datasets. Here is the mathematical definition:

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FN) + (TN + FP)}}$$

# References


[1] developers.google.com

[2] wikipedia.org/wiki/Precision_and_recall

[3] wikipedia.org/wiki/F-score

[4] wikipedia.org/wiki/Phi_coefficient