Classification Metrics

---

1. Confusion Matrix

Actual / Predicted	Positive (1)	Negative (0)
Positive (1)	TP (True Positive)	FN (False Negative)
Negative (0)	FP (False Positive)	TN (True Negative)

Key terms:

- TP = Correctly predicted positive
- TN = Correctly predicted negative
- FP = Predicted positive but actually negative
- FN = Predicted negative but actually positive

---

2. Accuracy

Measures overall correctness of the model.

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

✅ Works well when classes are balanced.

---

3. Precision (Positive Predictive Value)

Out of all predicted positives, how many are truly positive?

Formula:
Precision = TP / (TP + FP)

High precision → few false positives.

---

4. Recall (Sensitivity / True Positive Rate)

Out of all actual positives, how many did we correctly identify?

Formula:
Recall = TP / (TP + FN)

High recall → few false negatives.

---

5. F1-Score

Harmonic mean of Precision and Recall.

Formula:
F1 = 2 × (Precision × Recall) / (Precision + Recall)

Balances precision and recall.

---

6. Specificity (True Negative Rate)

Out of all actual negatives, how many did we correctly predict?

Formula:
Specificity = TN / (TN + FP)

---

7. False Positive Rate (FPR)

Proportion of negatives incorrectly predicted as positive.

Formula:
FPR = FP / (FP + TN)

---

8. False Negative Rate (FNR)

Proportion of positives incorrectly predicted as negative.

Formula:
FNR = FN / (FN + TP)

---

9. ROC Curve (Receiver Operating Characteristic)

- Plots TPR (Recall) vs FPR at various thresholds.
- Shows trade-off between sensitivity and specificity.
- Ideal model curve → close to top-left corner.

---

10. AUC (Area Under the ROC Curve)

Measures model's ability to distinguish classes.

Range: 0 to 1
Closer to 1 → better classifier.
0.5 → random guessing.

---

11. Precision–Recall (PR) Curve

Plots Precision vs Recall at different probability thresholds.
Useful for **imbalanced data**.

---

12. Log Loss (Cross-Entropy Loss)

Used in probabilistic classifiers (e.g., Logistic Regression).

Formula:
Log Loss = $-(1/N) \times \Sigma [y_i \times \log(p_i) + (1 - y_i) \times \log(1 - p_i)]$

Lower log loss = better model.

---

13. Matthews Correlation Coefficient (MCC)

Balanced metric for binary or imbalanced data.

Formula:
$MCC = (TP \times TN - FP \times FN) / \sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}$

Range:

- +1 → perfect
- 0 → random
- -1 → completely wrong

---

14. Cohen's Kappa

Measures agreement between predicted and actual labels, adjusted for chance.

Formula:
$Kappa = (p_o - p_e) / (1 - p_e)$

where
p_o = observed accuracy
p_e = expected accuracy (by chance)

---

15. Balanced Accuracy

Average of Sensitivity and Specificity.

Formula:
Balanced Accuracy = (Sensitivity + Specificity) / 2

---

16. Hamming Loss

Fraction of incorrectly predicted labels (for multi-label classification).

Formula:
$Hamming\ Loss = (1/N) \times \Sigma\ xor(y_i, \hat{y}_i)$

Lower value → better performance.

---

17. Jaccard Index (Intersection over Union)

Measures similarity between predicted and true sets.

Formula:
Jaccard = TP / (TP + FP + FN)

Used in **multi-label** and **image segmentation** tasks.

---

18. Macro, Micro, and Weighted Averages (for Multi-class)

Type	Description	Formula
Micro Average	Aggregates all classes before computing metric.	Compute global TP, FP, FN → then Precision/Recall/F1.
Macro Average	Simple mean of metrics for each class.	$(\text{Metric}_1 + \text{Metric}_2 + \ldots + \text{Metric}_n) / n$
Weighted Average	Weighted by number of samples per class.	$\Sigma (w_i \times \text{Metric}_i)$

---

19. Fβ-Score

Generalized F1-score; β controls importance of recall vs precision.

Formula:
$F\beta = (1 + \beta^2) \times (\text{Precision} \times \text{Recall}) / ((\beta^2 \times \text{Precision}) + \text{Recall})$

- $\beta > 1$ → focus on recall
- $\beta < 1$ → focus on precision

---

20. Top-K Accuracy

Used in multi-class problems (e.g., deep learning).

Formula:
Top-K Accuracy = (Number of samples where true label ∈ Top K predictions) / (Total samples)