Classification Metrics Overview

To calculate all probable measurement metrics for classification models, we consider the **actual (ground truth) values** and **predicted values**. Below are the key metrics:

1. Confusion Matrix Elements

- **True Positives (TP):** Correctly predicted positive cases.
- **False Positives (FP):** Incorrectly predicted positive cases (Type I Error).
- **True Negatives (TN):** Correctly predicted negative cases.
- **False Negatives (FN):** Incorrectly predicted negative cases (Type II Error).

2. Performance Metrics

(a) Accuracy

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

- Measures overall correctness.
- Works well for balanced datasets.

(b) Precision (Positive Predictive Value)

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

- Measures how many predicted positive cases are actually correct.
- Useful when false positives are costly (e.g., medical diagnosis).

(c) Recall (Sensitivity or True Positive Rate)

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

- Measures how well the model identifies actual positive cases.
- Important when false negatives are costly (e.g., fraud detection).

(d) F1-Score

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

- Harmonic mean of Precision and Recall.
- Balances both metrics.

(e) Specificity (True Negative Rate)

$$Specificity = \frac{TN}{TN + FP}$$

- Measures the ability to correctly identify negative cases.
- Important in medical tests (ensuring non-diseased cases are correctly classified).

(f) False Positive Rate (FPR)

$$FPR = \frac{FP}{FP + TN}$$

- The proportion of actual negatives incorrectly classified as positives.

(g) False Negative Rate (FNR)

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

- The proportion of actual positives incorrectly classified as negatives.

(h) Matthews Correlation Coefficient (MCC)

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

- Measures correlation between actual and predicted values.
- Values range from -1 (bad) to +1 (perfect).

(i) Cohen's Kappa Score

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

where:

- P_o = Observed agreement.
- P_e = Expected agreement by chance.

(j) Balanced Accuracy

$$BalancedAccuracy = \frac{Sensitivity + Specificity}{2}$$

- Useful for imbalanced datasets.

(k) ROC-AUC Score (Area Under the Receiver Operating Characteristic Curve)

- Measures the ability of the model to distinguish between classes.
- AUC = 1 means perfect classification; 0.5 means random guessing.

(l) PR-AUC Score (Area Under the Precision-Recall Curve)

- Useful for imbalanced datasets where precision and recall are critical.

Example of Confusion Matrix and finding all metrics.

Let's assume the following values from a **confusion matrix**:

	Actual Positive	**Actual Negative**
Predicted Positive	$TP = 50$	$FP = 10$
Predicted Negative	$FN = 5$	$TN = 35$

Now, let's compute all classification metrics.

---

1. Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{50 + 35}{50 + 35 + 10 + 5} = \frac{85}{100} = 0.85$$

(85%)

---

2. Precision (Positive Predictive Value)

$$Precision = \frac{TP}{TP + FP} = \frac{50}{50 + 10} = \frac{50}{60} = 0.8333$$

(83.33%)

---

3. Recall (Sensitivity / True Positive Rate)

$$Recall = \frac{TP}{TP + FN} = \frac{50}{50 + 5} = \frac{50}{55} = 0.9091$$

(90.91%)

---

4. F1-Score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0.8333 \times 0.9091}{0.8333 + 0.9091}$$
$$= 2 \times \frac{0.7575}{1.7424} = 2 \times 0.4347 = 0.8696$$

(86.96%)

---

5. Specificity (True Negative Rate)

$$Specificity = \frac{TN}{TN + FP} = \frac{35}{35 + 10} = \frac{35}{45} = 0.7778$$

(77.78%)

---

6. False Positive Rate (FPR)

$$FPR = \frac{FP}{FP + TN} = \frac{10}{10 + 35} = \frac{10}{45} = 0.2222$$

(22.22%)

---

7. False Negative Rate (FNR)

$$FNR = \frac{FN}{FN + TP} = \frac{5}{5 + 50} = \frac{5}{55} = 0.0909$$

(9.09%)

---

8. Matthews Correlation Coefficient (MCC)

$$MCC = \frac{(TP \times TN - FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$
$$= \frac{(50 \times 35 - 10 \times 5)}{\sqrt{(50 + 10)(50 + 5)(35 + 10)(35 + 5)}}$$
$$= \frac{(1750 - 50)}{\sqrt{60 \times 55 \times 45 \times 40}}$$
$$= \frac{1700}{\sqrt{5940000}} = \frac{1700}{2437.18} = 0.6975$$

(69.75%)

---

9. Cohen’s Kappa Score

First, calculate **Observed Agreement (Po)**:

$$P_o = \frac{TP + TN}{TP + TN + FP + FN} = \frac{50 + 35}{100} = 0.85$$

Calculate **Expected Agreement (Pe)**:

$$P_e = \left(\frac{(TP + FP) \times (TP + FN)}{N^2}\right) + \left(\frac{(TN + FN) \times (TN + FP)}{N^2}\right)$$

$$= \left(\frac{(50 + 10) \times (50 + 5)}{100^2}\right) + \left(\frac{(35 + 5) \times (35 + 10)}{100^2}\right)$$

$$= \left(\frac{60 \times 55}{10000}\right) + \left(\frac{40 \times 45}{10000}\right)$$

$$= \left(\frac{3300}{10000}\right) + \left(\frac{1800}{10000}\right) = 0.33 + 0.18 = 0.51$$

$$\kappa = \frac{P_o - P_e}{1 - P_e} = \frac{0.85 - 0.51}{1 - 0.51} = \frac{0.34}{0.49} = 0.6939$$

(69.39%)

---

10. Balanced Accuracy

$$BalancedAccuracy = \frac{Sensitivity + Specificity}{2} = \frac{0.9091 + 0.7778}{2}$$

$$= \frac{1.6869}{2} = 0.8434$$

(84.34%)

Summary of Metrics

Metric	Value (%)
Accuracy	85.00
Precision	83.33
Recall (Sensitivity)	90.91
F1-Score	86.96
Specificity	77.78
False Positive Rate	22.22
False Negative Rate	9.09
Matthews Correlation Coefficient (MCC)	69.75
Cohen’s Kappa Score	69.39
Balanced Accuracy	84.34