**🔷 What is a Confusion Matrix?**

A **Confusion Matrix** is a performance evaluation tool used for **classification problems**. It helps visualize the performance of a model by comparing the predicted labels with the actual labels.

It’s a **2x2 table** for binary classification:

|  |  |  |
| --- | --- | --- |
|  | **Predicted Positive** | **Predicted Negative** |
| **Actual Positive** | True Positive (TP) | False Negative (FN) |
| **Actual Negative** | False Positive (FP) | True Negative (TN) |

**✅ Terminologies:**

* **True Positive (TP)**: Model correctly predicted **Positive**.
* **False Positive (FP)**: Model incorrectly predicted **Positive** (Type I Error).
* **True Negative (TN)**: Model correctly predicted **Negative**.
* **False Negative (FN)**: Model incorrectly predicted **Negative** (Type II Error).

**📊 Example**

Let's say we are testing a medical test for detecting a disease:

* 100 people tested.
* 60 have the disease (Positive), 40 don’t (Negative).
* Model predicts:
  + 55 diseased people as Positive → **TP = 55**
  + 5 diseased people as Negative → **FN = 5**
  + 10 healthy people as Positive → **FP = 10**
  + 30 healthy people as Negative → **TN = 30**

**🧮 Confusion Matrix:**

|  | **Predicted Positive** | **Predicted Negative** |
| --- | --- | --- |
| **Actual Positive** | 55 (TP) | 5 (FN) |
| **Actual Negative** | 10 (FP) | 30 (TN) |

**📌 Performance Metrics & Their Formulae**

**1️⃣ Accuracy**

👉 Measures overall correctness.

**Formula:**

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

Using example:  
Accuracy = (55 + 30) / 100 = 85%

✅ Accuracy is useful when classes are balanced.

**2️⃣ Precision (Positive Predictive Value)**

👉 Out of all predicted positives, how many were correct?

**Formula:**

Precision = TP / (TP + FP)

**Using example:**

Precision = 55 / (55 + 10) ≈ 0.846

✅ High precision means low false positives.

**3️⃣ Recall (Sensitivity / True Positive Rate)**

👉 Out of all actual positives, how many did we catch?

**Formula:**

Recall = TP / (TP + FN)

**Using example:**

Recall = 55 / (55 + 5) ≈ 0.917

✅ High recall means low false negatives (good for medical diagnosis).

**4️⃣ F1 Score**

👉 Harmonic mean of Precision and Recall.  
Used when there’s an **imbalance** between Precision and Recall.

**Formula:**

F1 = 2 \* (Precision \* Recall) / (Precision + Recall)

**Using example:**

F1 = 2 \* (0.846 \* 0.917) / (0.846 + 0.917) ≈ 0.88

✅ Best when we want a **balance between Precision & Recall**.

**🔁 Relationships Between Metrics**

|  |  |  |
| --- | --- | --- |
| **Metric** | **Depends On** | **Indicates** |
| Accuracy | TP, TN, FP, FN | Overall correctness |
| Precision | TP, FP | How many selected items are relevant |
| Recall | TP, FN | How many relevant items are selected |
| F1 Score | Precision, Recall | Trade-off between precision and recall |

**⚠️ When Accuracy is Misleading**

Consider an imbalanced dataset:

* 990 Negative, 10 Positive
* Model predicts all as Negative

Confusion Matrix:

|  |  |  |
| --- | --- | --- |
|  | **Pred. Pos** | **Pred. Neg** |
| Actual Positive | 0 | 10 |
| Actual Negative | 0 | 990 |

* **Accuracy** = 990/1000 = 99% ✅
* But **Recall** = 0 → Model misses all positive cases ❌
* Use **F1 Score or Recall** instead.

**📌 Summary Table**

|  |  |  |
| --- | --- | --- |
| **Metric** | **Formula** | **Good For** |
| Accuracy | (TP + TN) / (TP + TN + FP + FN) | Balanced datasets |
| Precision | TP / (TP + FP) | Low false positive requirement |
| Recall | TP / (TP + FN) | Low false negative requirement |
| F1 Score | 2 \* (P \* R) / (P + R) | Balance between P and R |

**✅ Tip to Remember:**

* **Precision** = *What proportion of predicted Positives is correct?*
* **Recall** = *What proportion of actual Positives was identified?*
* **F1** = *Balance between Precision and Recall.*