**1. Accuracy (Exactitude)**

**Définition :** Pourcentage de prédictions correctes (positives **et** négatives) par rapport au total.

**Formule :**  
Accuracy = (TP + TN) / (TP + TN + FP + FN)

**TP** : vrai positif (bonne détection d’une anomalie)  
**TN** : vrai négatif (bonne détection d’un cas normal)  
**FP** : faux positif (fausse alerte)  
**FN** : faux négatif (anomalie non détectée)

**Mais attention** : si les anomalies sont très rares, l’accuracy peut être **trompeuse**. Ex. : 99% de "normaux" → prédire tout normal = 99% de bonnes réponses, mais ça ne détecte **aucune anomalie**.

**2. Precision (Précision)**

Parmi tous les cas que le modèle a détectés comme anomalies, combien étaient vraiment des anomalies ?

**Formule :**  
Precision = TP / (TP + FP)

Plus la **précision est haute**, moins il y a de **faux positifs**.

**3. Recall (Rappel ou Sensibilité)**

Parmi toutes les vraies anomalies, combien le modèle a réussi à détecter ?

**Formule :**  
Recall = TP / (TP + FN)

Plus le **recall est élevé**, moins il y a de **faux négatifs** (anomalies manquées).

**⚖️ 4. F1-score**

Moyenne harmonique entre **précision** et **recall**. Un bon compromis quand on veut à la fois éviter les fausses alertes **et** rater des anomalies.

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

* Un score F1 proche de 1 = excellent équilibre
* Si F1 est faible, soit tu as trop de faux positifs, soit trop de faux négatifs

**5. Robustesse (Robustness)**

Capacité du modèle à rester **performant** même dans des **conditions variées ou bruitées**.

Ce n’est pas une métrique unique, mais une **qualité générale** du modèle. On peut l’évaluer en :

* testant sur d’autres jeux de données
* introduisant du bruit dans les données
* vérifiant la stabilité des résultats face aux changements

**🧠 En résumé avec une image mentale :**

| **Situation** | **Exemple réel** | **Métrique clé** |
| --- | --- | --- |
| Beaucoup d’alertes justes ? | Tu détectes bien les anomalies | Précision haute |
| Tu rates peu d’anomalies ? | Tu n’en oublies pas | Rappel haut |
| Équilibre entre les deux ? | Bon détecteur global | F1-score |
| Peu d’erreurs globales ? | Peu de mauvaises prédictions | Accuracy |
| Résiste aux perturbations ? | Résultats cohérents même si le contexte change | Robustesse |