

Introduction à la classification supervisée

Matrice de confusion, Accuracy, Précision, Rappel, F1-score

K. Kadri

- ① Métriques en Machine Learning
 - Perspective mathématique
 - Métriques en supervisé
 - Métriques en non-supervisé
- ② Contexte : classification supervisée
- ③ Matrice de confusion
- ④ Accuracy
- ⑤ Précision et Rappel
- ⑥ F1-score
- ⑦ Au-delà du binaire
- ⑧ Résumé

Pourquoi les métriques sont essentielles ?

- Une métrique = une **fonction mathématique** qui mesure la qualité d'un modèle.
- Elles guident :
 - la sélection de modèles,
 - le tuning d'hyperparamètres,
 - l'arrêt du training,
 - l'évaluation finale.

Idée clé

Pas de bonne IA sans bonnes métriques. Une métrique détermine ce que “réussir” veut dire pour un modèle.

Une métrique est une fonction :

$$M : (y_{\text{réel}}, y_{\text{prédit}}) \mapsto \mathbb{R}$$

qui doit être :

- **cohérente** : 2 prédictions proches donnent 2 scores proches ;
- **interprétable** : bornée ou normalisée si possible ;
- **optimisable** : adaptée aux algorithmes d'apprentissage.

Exemple

En classification :

$$\text{Accuracy} = \frac{1}{n} \sum_{i=1}^n 1(y_i = \hat{y}_i)$$

- **Régression**

- MAE :

$$\text{MAE} = \frac{1}{n} \sum |y_i - \hat{y}_i|$$

- MSE :

$$\text{MSE} = \frac{1}{n} \sum (y_i - \hat{y}_i)^2$$

- R^2 :

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

- **Classification**

- Accuracy
- Précision, Rappel
- F1-score
- AUC-ROC

Quand choisir quelle métrique ?

- Classes équilibrées → **Accuracy**
- Faux positifs coûteux → **Précision**
- Faux négatifs graves → **Rappel**
- Déséquilibre fort → **F1** ou **AUC-ROC**

Message clé

Une métrique = un ****choix métier****, pas seulement mathématique.

Non-supervisé : pas de vérité terrain !

- Pas de $y_{\text{réel}}$ \rightarrow les métriques mesurent des **cohérences internes**.
- On mesure :
 - compacité des clusters,
 - séparation des clusters,
 - densité locale,
 - stabilité structurelle.

- Silhouette score :

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

- $a(i)$ = distance intra-cluster
- $b(i)$ = distance au cluster le plus proche
- **Davies–Bouldin index** Plus il est petit, meilleur est le clustering.
- **Calinski–Harabasz** Plus grand = mieux.

Intuition

Un bon cluster est ****compact**** et ****bien séparé**** des autres.

- Distance-based :

$$D(x) = \|x - \mu\|$$

- Density-based (LOF)

$$\text{LOF}(x) = \frac{\text{densité locale moyenne}}{\text{densité locale de } x}$$

Usage pratique

Une anomalie = point peu dense ou isolé. La métrique mesure cet éloignement.

Classification supervisée : idée générale

- Données étiquetées : **caractéristiques** (features) + **label** (classe).
- Objectif : apprendre une fonction $f(x) \rightarrow y$ qui prédit la bonne classe.
- Exemples :
 - Diagnostic médical : malade / sain.
 - Email : spam / non spam.
 - Qualité produit : conforme / non conforme.

Idée clé

En classification, on ne s'intéresse pas seulement au **taux d'erreur global**, mais à **comment** le modèle se trompe (type d'erreurs).

Matrice de confusion : binaire

- Problème binaire : **Positif / Négatif**.
- La matrice de confusion compare :
 - Les **vraies étiquettes**,
 - Aux **prédictions** du modèle.
- **TP** : True Positive (vrai positif)
- **TN** : True Negative (vrai négatif)
- **FP** : False Positive (faux positif, « faux alarmes »)
- **FN** : False Negative (faux négatif, « ratés »)

	Prédit	
	Positif	Négatif
Réel Positif	TP	FN
Réel Négatif	FP	TN

Exemple simple

Supposons 100 exemples, avec la matrice de confusion :

	Prédit	
	Positif	Négatif
Réel Positif	40	10
Réel Négatif	5	45

- $TP = 40$, $FN = 10$, $FP = 5$, $TN = 45$.
- $Total = 40 + 10 + 5 + 45 = 100$.

Question pour le TD

À partir de cette matrice, calculez : Accuracy, Précision, Rappel, F1-score.

Accuracy (taux de bonne classification)

Définition

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

- Proportion de prédictions **correctes** parmi tous les exemples.
- Utile quand :
 - Les classes sont **équilibrées**.
 - Les **coûts d'erreurs** FP et FN sont similaires.

Limite

Si les données sont très **déséquilibrées** (ex : 99% de négatifs), un modèle trivial peut avoir une accuracy élevée sans être utile.

Définition

$$\text{Précision} = \frac{TP}{TP + FP}$$

- Parmi tous les exemples **prédits positifs**, combien sont vraiment positifs ?
- Mesure la qualité des **positifs prédits**.

Quand la privilégier ?

Quand les **faux positifs** sont coûteux. Exemple : envoyer un email marketing à un client non intéressé, alerter à tort.

Définition

$$\text{Rappel} = \frac{TP}{TP + FN}$$

- Parmi tous les **vrais positifs**, combien sont détectés ?
- Mesure la capacité à **ne pas rater** les positifs.

Quand le privilégier ?

Quand les **faux négatifs** sont très graves. Exemple : cancer non détecté, fraude laissée passer.

F1-score : compromis précision / rappel

Définition

$$F1 = 2 \times \frac{\text{Précision} \times \text{Rappel}}{\text{Précision} + \text{Rappel}}$$

- Moyenne **harmonique** entre précision et rappel.
- Valeur élevée seulement si **les deux** sont élevées.

Intérêt

- Utile sur des données **déséquilibrées**.
- Bon indicateur global de la qualité de détection de la classe positive.

Exemple de calcul (avec la matrice précédente)

Rappel de la matrice :

$$TP = 40, \quad FN = 10, \quad FP = 5, \quad TN = 45.$$

- Accuracy :

$$Acc = \frac{40 + 45}{100} = 0,85$$

- Précision :

$$Prec = \frac{40}{40 + 5} = \frac{40}{45} \approx 0,89$$

- Rappel :

$$Rec = \frac{40}{40 + 10} = \frac{40}{50} = 0,8$$

- F1 :

$$F1 = 2 \times \frac{0,89 \times 0,8}{0,89 + 0,8} \approx 0,84$$

À proposer en exercice

Cas multi-classes (3 classes ou plus)

- La matrice de confusion devient une **grille** $K \times K$.
- On peut calculer précision, rappel, F1 :
 - **Par classe** (one-vs-rest),
 - Puis faire une moyenne :
 - Macro-average (moyenne simple des classes),
 - Weighted-average (pondérée par la taille des classes).

Lien avec le TD

Sur un dataset multi-classes (ex : *wine*, *digits*), demander aux étudiants d'afficher la matrice de confusion et les F1 *par classe*.

- **Matrice de confusion** : base pour tout analyser.
- **Accuracy** : vision globale, attention aux classes déséquilibrées.
- **Précision** : qualité des positifs prédits (FP coûteux).
- **Rappel** : capacité à capturer les vrais positifs (FN coûteux).
- **F1-score** : compromis précision / rappel.

Message clé pour le TD

Toujours choisir la ou les métriques en fonction du **contexte métier** (coût des faux positifs / faux négatifs, déséquilibre des classes, etc.).