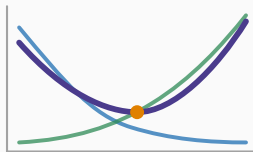


Apprentissage statistique

Chapitre 7 : Évaluer un classifieur : matrice de confusion, ROC, AUC



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

La matrice de confusion

Le piège de l'exactitude

La courbe ROC

Ce qu'il faut retenir

La matrice de confusion

Quatre nombres

Les quatre cases

	Réel : défaut	Réel : bon
Prévu : défaut	vrais positifs VP	faux positifs FP
Prévu : bon	faux négatifs FN	vrais négatifs VN

Notre modèle sur les mille dossiers

	Réel : défaut	Réel : bon	Total
Prévu : défaut	60	90	150
Prévu : bon	40	810	850
Total	100	900	1 000

Les mesures qu'on en tire

$$VP + VN = 870$$

Le piège de l'exactitude

Un concurrent redoutable

Considérons un modèle qui prédit « **bon** » pour tout le monde, sans regarder aucune variable.

$$\text{exactitude} = \frac{900}{1\,000} = 0,90$$

Il fait mieux que notre modèle — 0,90 contre 0,87 — tout en étant strictement inutile : son rappel est nul, il ne détecte **aucun** défaut.

La leçon

Avec des classes déséquilibrées, l'exactitude est dominée par la classe majoritaire. Elle récompense l'inaction.

Une exactitude élevée n'établit rien tant qu'on ne l'a pas comparée à celle du classifieur trivial. C'est le premier réflexe à acquérir, et l'erreur la plus répandue chez les débutants.

Définition

Abaissier le seuil augmente le **rappel** — on détecte plus de défauts — et dégrade la **précision** — on se trompe plus souvent en accusant.

Le **F1** les combine par leur moyenne harmonique :

$$F_1 = 2 \times \frac{\text{précision} \times \text{rappel}}{\text{précision} + \text{rappel}} = 2 \times \frac{0,40 \times 0,60}{1,00} = 0,48$$

Lequel privilégier dépend du coût des erreurs

Application en gestion

Contexte	Priorité
Dépistage d'une maladie grave	le rappel — ne rien manquer
Filtre anti-spam	la précision — ne pas bloquer un message légitime
Octroi de crédit	selon le coût relatif du défaut et de la marge perdue

Aucune mesure unique ne convient à tous les problèmes. Le choix se fait sur l'économie du cas.

La courbe ROC

Courbe ROC

La **courbe ROC** trace, pour tous les seuils possibles, le **taux de vrais positifs** — le rappel — contre le **taux de faux positifs**, soit $1 - \text{spécificité}$.

Sur notre modèle et à ce seuil, le point vaut $(0,10; 0,60)$.

Comment la lire

Position	Signification
Coin supérieur gauche	classifieur parfait
Diagonale	classifieur aléatoire
Sous la diagonale	pire que le hasard — inverser les prédictions

Définition

L'AUC est l'aire sous la courbe ROC.

AUC	Lecture
0,5	aléatoire
0,7 à 0,8	acceptable
> 0,9	excellent — ou suspect

Intuition

L'AUC est la **probabilité** que le modèle attribue un score plus élevé à un défaillant tiré au hasard qu'à un bon client tiré au hasard.

Cette interprétation est précieuse : elle ne dépend d'**aucun seuil**, et elle est insensible au déséquilibre des classes. C'est ce qui en fait la mesure de référence pour comparer des modèles.

Application en gestion

Une AUC supérieure à 0,95 sur un problème de crédit doit déclencher une vérification, non une célébration.

Les causes habituelles sont la **fuite de données** du chapitre 1 — une variable qui contient l'information à prévoir, comme le nombre de relances ou la date de clôture du dossier.

Le premier réflexe devant un résultat excellent est de chercher la fuite.

Ce qu'il faut retenir

Cinq points

1. La **matrice de confusion** résume tout : ici 60, 40, 90, 810.
2. L'**exactitude** de 0,87 est battue par le classifieur trivial à 0,90 — elle ne prouve rien seule.
3. **Précision** et **rappel** s'opposent; le seuil arbitre entre eux; $F_1 = 0,48$ les combine.
4. La **courbe ROC** balaie tous les seuils; ici le point (0,10 ; 0,60).
5. L'**AUC** est la probabilité de bien ordonner deux dossiers; une valeur trop élevée signale une fuite.

La suite

Nous savons construire et évaluer des modèles linéaires. Les cinq chapitres suivants abandonnent la linéarité, en commençant par l'arbre de décision.