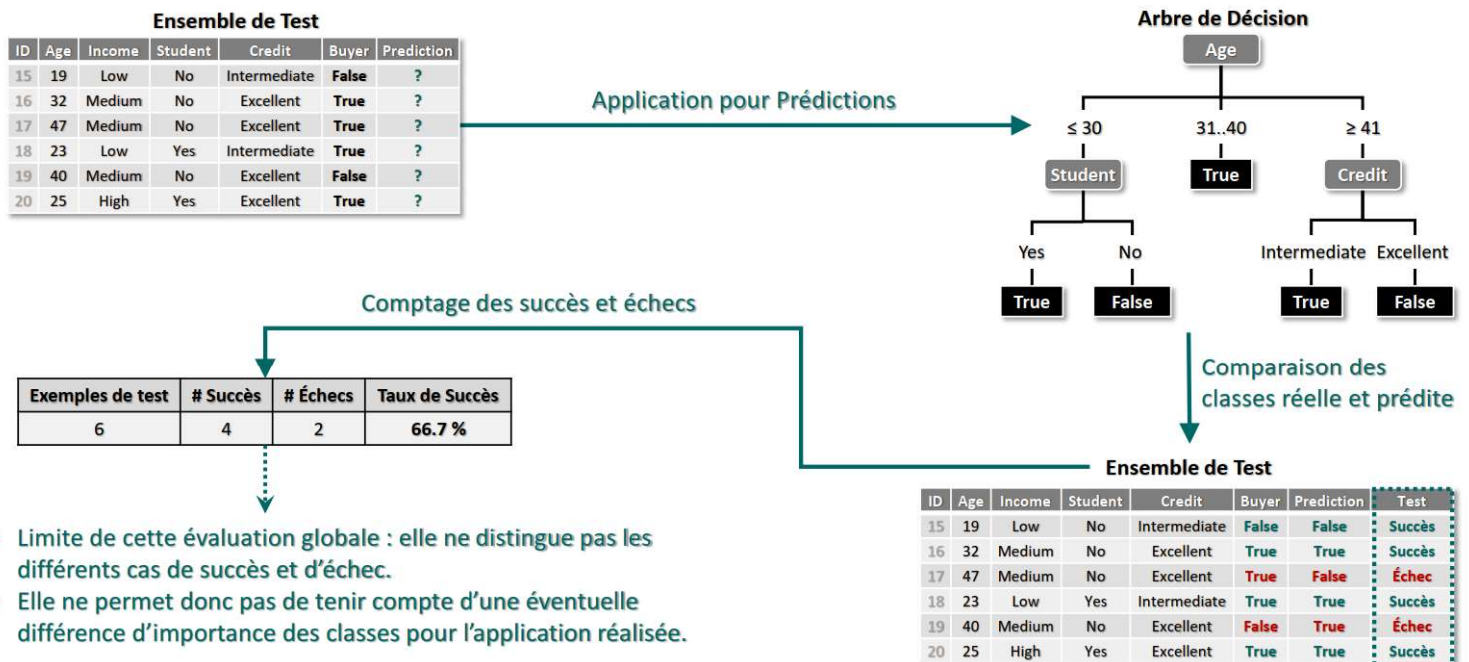


Matrices de Confusion

Nicolas PASQUIER
 Université Côte d'Azur
 Département Informatique
 Laboratoire I3S (UMR-7271 UCA/CNRS)
<http://www.i3s.unice.fr/~pasquier>



Rappel : Évaluation des Classifieurs par Taux de Succès



Intérêt de la Matrice de Confusion

- Outil d'évaluation qui distingue les différents types de succès et d'échec pour chaque classe
 - Détaille le nombre de prédictions correctes et incorrectes pour chaque classe prédite et chaque classe réelle (classe de l'exemple de test)
 - Permet de quantifier numériquement l'importance des différents types de succès et d'erreurs
- Calculée comme une matrice de contingence dans laquelle
 - Les lignes représentent la classe réelle des exemples de test
 - Les colonnes représentent la classe prédite pour les exemples de test
- Outil central pour l'évaluation et la comparaison des performances des classifieurs
 - L'interprétation des matrices de confusions permet de comparer la qualité de la discrimination des classes par les classifieurs
 - Permet le calcul automatique de mesures de qualités des prédictions

Calcul de la Matrice de Confusion

- Exemple : évaluation de l'arbre de décision de prédiction d'appétence (Buyer)

Ensemble de Test

ID	Age	Income	Student	Credit	Buyer	Prediction	Test
15	19	Low	No	Intermediate	False	False	Succès
16	32	Medium	No	Excellent	True	True	Succès
17	47	Medium	No	Excellent	True	False	Échec
18	23	Low	Yes	Intermediate	True	True	Succès
19	40	Medium	No	Excellent	False	True	Échec
20	25	High	Yes	Excellent	True	True	Succès

Comptage des cas de succès et d'échecs pour chaque classe réelle et prédiction faite :

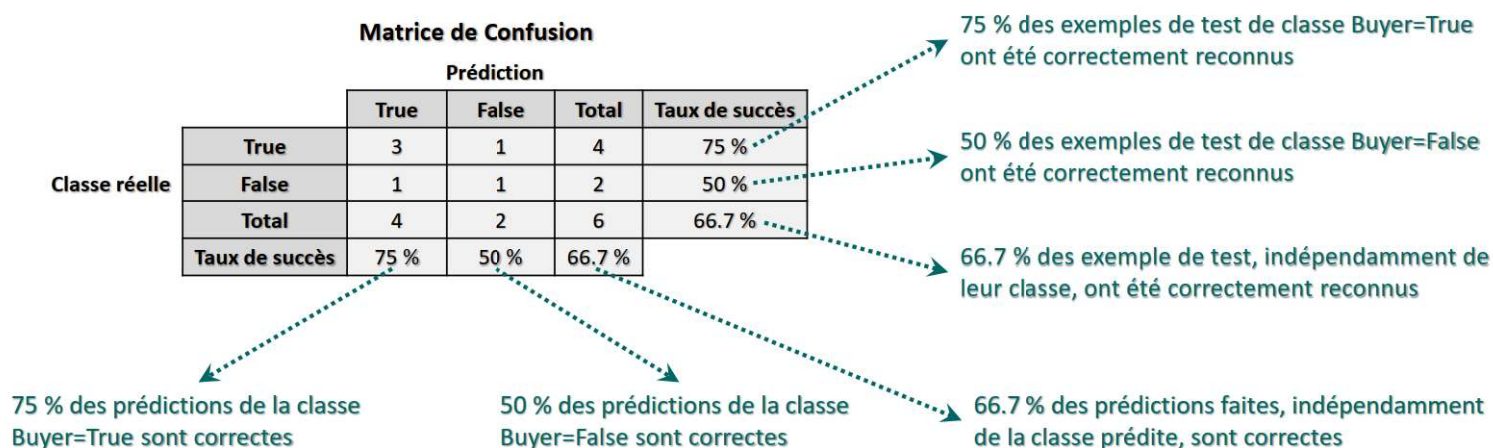
- 3 prédictions de la classe Buyer=True sont correctes (exemples de test de classe réelle Buyer=True)
- 1 prédiction de la classe Buyer=False est correcte (exemple de test de classe réelle Buyer=False)
- 1 prédiction de la classe Buyer=True est incorrecte (exemple de test de classe réelle Buyer=False)
- 1 prédiction de la classe Buyer=False est incorrecte (exemple de test de classe réelle Buyer=True)

Matrice de Confusion

		Prédiction		
		True	False	Total
Classe réelle	True	3	1	4
	False	1	1	2
Total		4	2	6

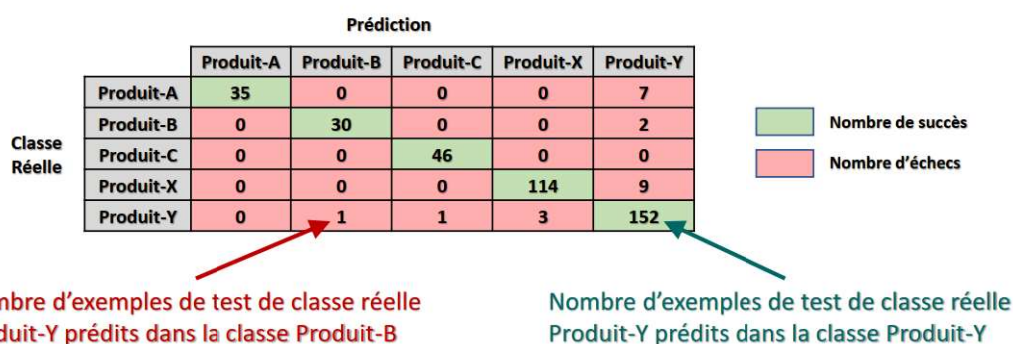
Interprétation de la Matrice de Confusion

- Les valeurs permettent de distinguer quatre situations et obtenir des indicateurs globaux (i.e. sans distinction des classes) sur la fiabilité du classifieur
- Exemple : interprétation en terme de taux de succès



Matrices de Confusion : Classification Multi-Classes

- Classification binaire : deux classes à prédire, e.g. Buyer \in {True, False}
- Classification multi-classes : plus de deux classes à prédire, e.g. Médicament \in {Produit-A, Produit-B, Produit-C, Produit-X, Produit-Y}
 - Autant de lignes et colonnes que de classes à prédire
 - Diagonale principale : nombres de succès
 - Autres cellules : nombres d'échecs



Classes Positive et Négative

- Si les objectifs de l'application donnent plus d'importance à une classe, celle-ci est désignée classe positive et l'autre classe désignée classe négative
- Exemple : applications de prédiction d'appétence Buyer=True ou Buyer=False
 - On privilégie la classe des acheteurs potentiels (i.e. prédictions Buyer=True) auxquels on enverra un message publicitaire par rapport à celle des acheteurs improbables (i.e. prédictions Buyer=False)
 - La classe Buyer=True est donc la classe positive et la classe Buyer=False la classe négative
 - Interprétation : on souhaite minimiser les erreurs de prédictions Buyer=True quitte à augmenter le nombre d'erreurs de prédictions Buyer=False

Matrice de Confusion

	Prédiction		
	Positive	Négative	Total
Classe réelle			
Positive	3	1	4
Négative	1	1	2
Total	4	2	6

Classes Positive et Négative

- Parmi les prédictions positives et négatives, on distingue les vrais (succès) et les faux (échecs)

Vrais Positifs (VP) : nombre de prédictions positives correctes.

⇒ Nombre d'exemples de test positifs prédits positifs

Faux Négatifs (FN) : nombre de prédictions négatives incorrectes.

⇒ Nombre d'exemples de test positifs prédits négatifs

Matrice de Confusion

	Prédiction		
	Positive	Négative	Total
Classe réelle			
Positive	3	1	4
Négative	1	1	2
Total	4	2	6

Faux Positifs (FP) : nombre de prédictions positives incorrectes.

⇒ Nombre d'exemples de test négatifs prédits positifs

Vrais Négatifs (VN) : nombre de prédictions négatives correctes.

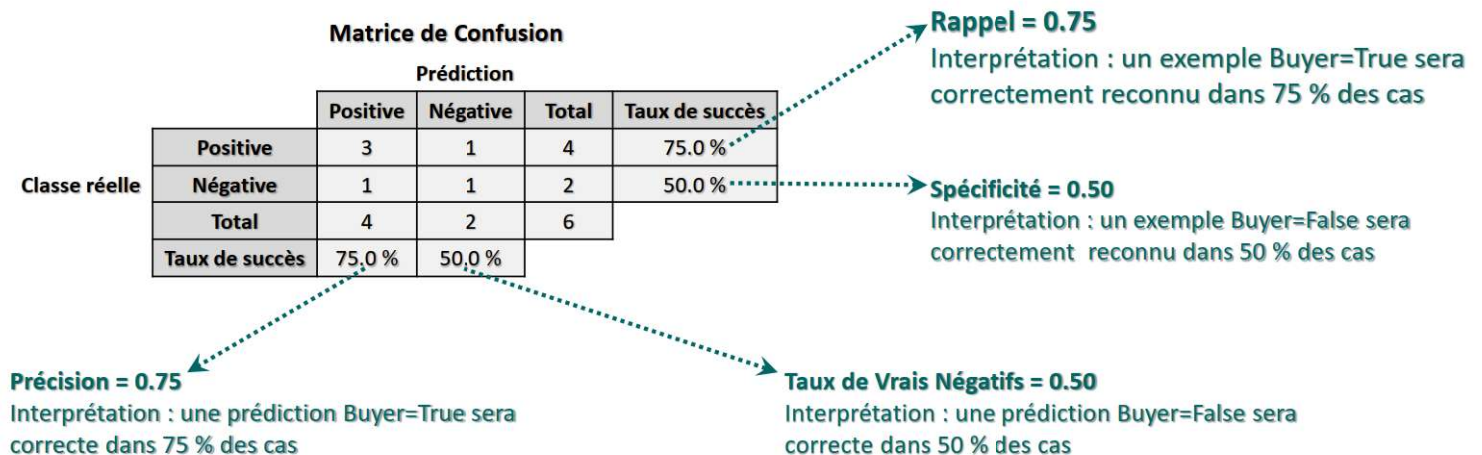
⇒ Nombre d'exemples de test négatifs prédits négatifs

Mesures d'Évaluation des Succès et Échecs

- Les valeurs VP, FP, VN et FN permettent de calculer des mesures d'évaluation
- Rappel (Sensibilité)
 - Calcul : $VP / (VP + FN)$
 - Proportion d'exemples de test positifs correctement reconnus
- Spécificité
 - Calcul : $VN / (VN + FP)$
 - Proportion d'exemples de test négatifs correctement reconnus
- Précision
 - Calcul : $VP / (VP + FP)$
 - Proportion de prédictions positives correctes
- Taux de vrais négatifs
 - Calcul : $VN / (VN + FN)$
 - Proportion de prédictions négatives correctes

Interprétation des Mesures d'Évaluation

- Ces mesures permettent de sélectionner le classifieur qui maximise les succès soit pour les positifs (Rappel et Précision), soit pour les négatifs (Spécificité et Taux de Vrais Négatifs)

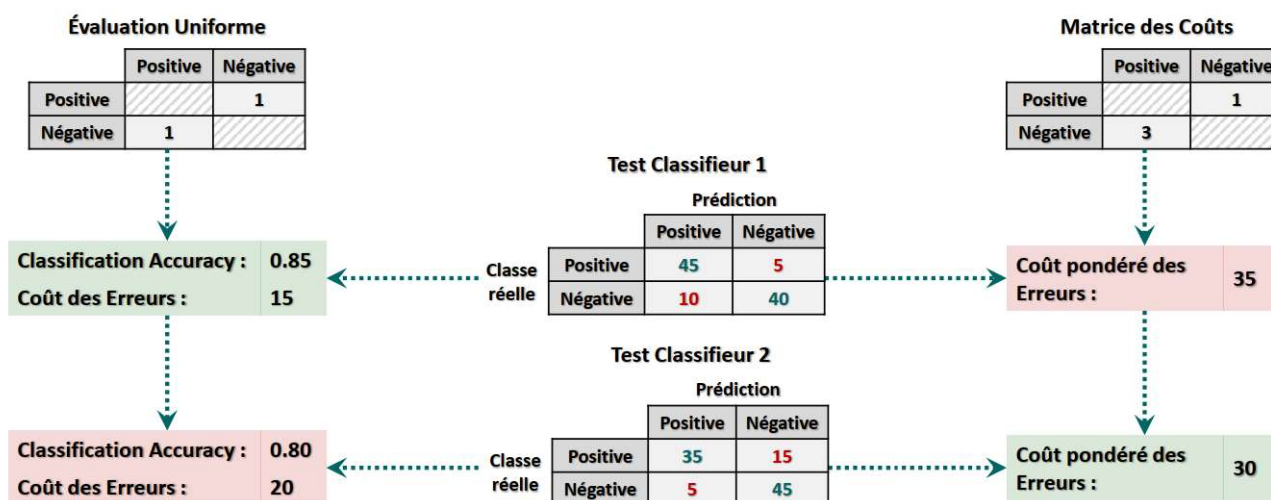


Exemple : Application de Prédiction d'Appétence

- Dans le cadre d'une campagne de marketing par message publicitaires
 - Les clients prospects prédits dans la classe Buyer=True recevront un message publicitaire (classe positive)
 - Les clients prospects prédits dans la classe Buyer=False seront ignorés (classe négative)
- Maximiser la Précision permet de maximiser les chances d'envoyer un message publicitaire utile (minimiser les coûts)
- Maximiser le Rappel permet de minimiser les chances de ne pas envoyer un message publicitaire utile (rater un client potentiel)
- Maximiser le Taux de Vrais Négatifs permet de minimiser les chances d'envoyer un message publicitaire inutile (éviter les « spams »)
- Maximiser la Spécificité permet de minimiser les chances de ne pas envoyer un message publicitaire inutile (ne pas détecter un client improbable)

Matrice des Coûts

- Permet de quantifier numériquement l'importance relative des différents types d'erreurs en attribuant un coût variable (poids) à chaque type d'erreur
- Exemple : coût des faux positifs triple par rapport aux faux négatifs



Références et Bibliographie

- Principales Librairies R
 - [heuristic](#) : calculs de mesures d'évaluation et heuristiques de sélection des classifieurs
 - [mlearning](#) : calcul de matrices de confusion et affichages graphiques d'exploration des résultats
 - [ROCR](#) : calculs de mesures d'évaluation et de représentation graphique pour l'évaluation des performances de classifieurs
- Bibliographie
 - Data Classification: Algorithms and Applications. Chapter 24 (Evaluation of Classification Methods). Charu C. Aggarwal. Chapman and Hall/CRC, 2014. ISBN 978-1-466-58674-1
 - Data Mining Applications with R. Yanchang Zhao & Yonghua Cen. Academic Press, 2014. ISBN 978-0-124-11511-8