

Apprentissage automatique

Évaluation

Amira Barhoumi

`amira.barhoumi@univ-grenoble-alpes.fr`

Année universitaire : 2025-2026

Apprendre à classifier

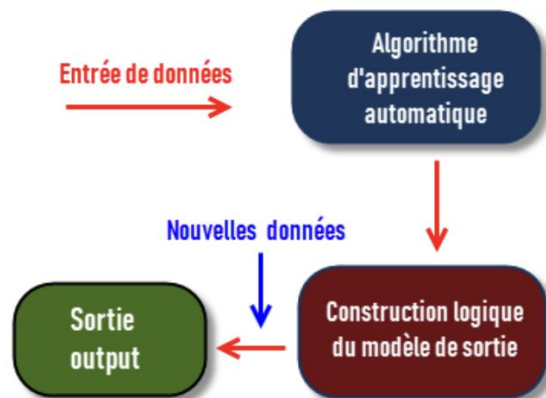
Un problème d'apprentissage fait intervenir trois composants essentiels:

- la tâche à résoudre
- les exemples (le corpus, le jeu de données)
- la mesure de performance (nombre d'erreurs, taux de bonne classification, etc)

Apprendre à classifier

Après la phase d'entraînement, la performance du modèle est mesurée par sa capacité à classifier correctement de nouveaux exemples

- Étapes d'apprentissage automatique :
 - Préparation des données
 - Construction du modèle
 - Évaluation et/ou analyse des résultats



- Étapes d'apprentissage automatique :
 - Préparation des données
 - Construction du modèle
 - Évaluation et/ou analyse des résultats

x1	x2	x3	y
1	10	100	alpha
2	20	200	beta

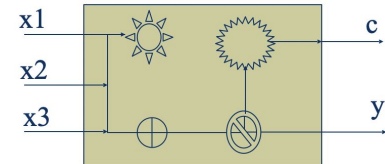
Entrée de données
→

Algorithme
d'apprentissage
automatique

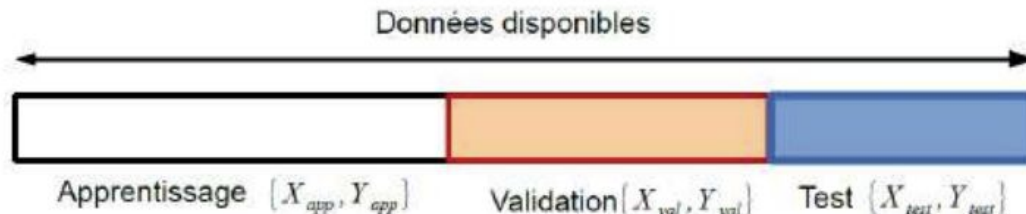
Nouvelles données
↓

Sortie
output

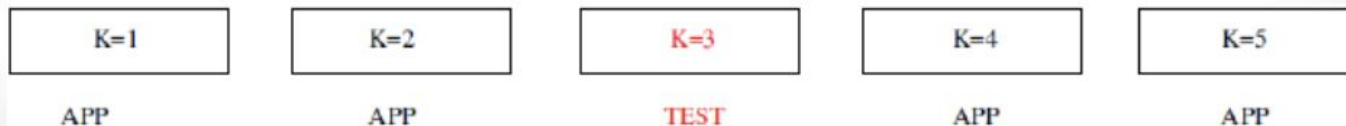
Construction logique
du modèle de sortie



- Validation classique :
 - Corpus d'apprentissage
 - Corpus de validation
 - Corpus de test
- Division du corpus en différents ensembles
 - 70% vs. 30% ou bien 80% vs. 20%
 - 70% vs. 10% vs. 20%



- Validation croisée (*cross-validation*):
 - Corpus modeste de dimension N
 - Estimation de l'erreur de généralisation par rééchantillonnage
 - Principe:
 - Séparer les N données en K ensembles de dimension égale
 - Pour chaque $k = 1, \dots, K$
 - Apprendre un modèle en utilisant les $K-1$ autres ensembles de données
 - Évaluer le modèle sur le k -ième ensemble
 - Calculer la moyenne des k estimations de performances obtenues pour obtenir la performance finale par validation croisée



- Erreur :
 - Erreur de prédiction :
 - évaluer le modèle sur le corpus d'apprentissage
 - taux de la mauvaise classification sur le corpus d'apprentissage
 - Erreur de généralisation
 - évaluer le modèle sur le corpus de test
 - taux de la mauvaise classification sur le corpus de test
 - Corpus de test

- Mesure de performance du modèle :
 - Y_i la classe référence (gold)
 - Y_i' la classe prédite par le modèle
 - Évaluer l'erreur $Y_i' - Y_i$
 - Dépendre de l'objectif de l'application visée
 - dans le domaine de la recherche d'information
 - précision
 - rappel
 - dans le domaine de la traduction automatique
 - Bleu
 - Rouge

- Mesure de performance du modèle
- Tableau de contingence (matrice de confusion)

matrice de confusion		prédiction	
		positif	négatif
référence	positif	TP	FN
	négatif	FP	TN

- Mesure de performance du modèle
- Taux de bonne classification

matrice de confusion		prédiction	
		positif	négatif
référence	positif	TP	FN
	négatif	FP	TN

TP rate = (# instances bien classées) / (# instances classées)

- Mesure de performance du modèle
- Taux de mauvaise classification
- Taux d'erreur

matrice de confusion		prédiction	
		positif	négatif
référence	positif	TP	FN
	négatif	FP	TN

Error rate = (# instances mal classées) / (# instances classées)

- Mesure de performance du modèle
- Précision : le pourcentage des

documents pertinents (correctement classés)

Calculer la précision par classe puis
calculer la moyenne des précisions
obtenues \Rightarrow macro-précision

matrice de confusion		prédiction	
		positif	négatif
référence	positif	TP	FN
	négatif	FP	TN

précision_{ci} = (# instances correctement classées c_i) / (# instances classées c_i)

- Mesure de performance du modèle
- Rappel : le pourcentage des

documents pertinents retrouvés

Calculer le rappel par classe puis
calculer la moyenne des rappels
obtenus \Rightarrow macro-rappel

matrice de confusion		prédiction	
		positif	négatif
référence	positif	TP	FN
	négatif	FP	TN

rappel_{ci} = (# instances correctement classées c_i) / (# instances réellement c_i)

- Mesure de performance du modèle
- F-mesure : combinaison du

rappel et de la précision

matrice de confusion		prédiction	
		positif	négatif
référence	positif	TP	FN
	négatif	FP	TN

F-mesure = $(1+\beta) * \text{rappel} * \text{précision} / (\beta^2 * (\text{rappel} + \text{précision}))$

Si $\beta = 1$, alors c'est la F1-mesure

- Exemple

Soit le corpus de test composé des objets obj1, ..., obj11.

La colonne hypothèse représente la prédiction d'un modèle d'apprentissage sur ce corpus de test. Le modèle n'arrive pas à prédire une classe pour l'objet obj11.

Dresser la matrice de confusion

Calculer les mesures suivantes :

- exactitude
- précision
- rappel
- F1-mesure

	référence	hypothèse
obj1	oui	oui
obj2	oui	oui
obj3	oui	non
obj4	non	oui
obj5	oui	oui
obj6	non	oui
obj7	non	non
obj8	oui	oui
obj9	non	oui
obj10	non	non
obj11	non	-

- Exemple

matrice de confusion		prédiction	
		oui	non
référence	oui	4	1
	non	3	2

$$\begin{aligned} Exactitude &= \frac{4 + 2}{10} \\ &= \frac{6}{10} \\ &= 0.6 \end{aligned}$$

- Exemple

$$\begin{aligned} Precision_{oui} &= \frac{4}{4+3} \\ &= \frac{4}{7} \\ &= 0.5714 \end{aligned}$$

$$\begin{aligned} Precision_{non} &= \frac{2}{2+1} \\ &= \frac{2}{3} \\ &= 0.6666 \end{aligned}$$

$$\begin{aligned} Precision &= \frac{0.5714 + 0.6666}{2} \\ &= 0.619 \end{aligned}$$

$$\begin{aligned} Rappel_{oui} &= \frac{4}{5} \\ &= 0.8 \end{aligned}$$

$$\begin{aligned} Rappel_{non} &= \frac{2}{6} \\ &= 0.3333 \end{aligned}$$

$$\begin{aligned} Rappel &= \frac{0.8 + 0.3333}{2} \\ &= 0.5666 \end{aligned}$$

$$\begin{aligned} F1 - mesure &= \frac{2 \times Rappel \times Precision}{Rappel + Precision} \\ &= \frac{2 \times 0.5666 \times 0.619}{0.5666 + 0.619} \\ &= 0.5916 \end{aligned}$$