

UNIVERSITÉ
PAUL VALÉRY
MONTPELLIER 3












EVALUATION DES MODELES

sandra.bringay@univ-montp3.fr

UNIVERSITÉ
PAUL VALÉRY
MONTPELLIER 3

CLASSIFICATION

- comestibles (C)
- non comestibles (N)
- toxiques (T)
- mortels (M)

 (C)	 (T)	 (C)	 (C)
 (M)	 (T)	 (M)	 (T)
 (C)	 (N)	...	 (M)












<http://www.univ-montp3.fr/miap/ens/info/>

2

UNIVERSITÉ
PAUL VALÉRY
MONTPELLIER 3

CLASSIFICATION

- comestibles (C)
- non comestibles (N)
- toxiques (T)
- mortels (M)

 (C)	 (T)	 (C)	 (C)
 (M)	 (T)	 (M)	 (T)
 (C)	 (N)	...	 (M)

Apprentissage

Test









<http://www.univ-montp3.fr/miap/ens/info/>

3

UNIVERSITÉ
PAUL VALÉRY
MONTPELLIER 3

ETAPE 1

Apprentissage :
Construction du modèle

 (C)	 (T)	 (C)
 (M)	 (T)	 (M)
 (C)	 (N)	...




<http://www.univ-montp3.fr/miap/ens/info/>

4

UNIVERSITÉ
PAUL VALÉRY
MONTPELLIER 3

ETAPE 2

- Vérifier l'étiquette proposée par le classifieur
- Calculer la performance du modèle

 (C)	✗ (M)
 (T)	✓
 (M)	✗ (N)

<http://www.univ-montp3.fr/miap/ens/info/>


5

UNIVERSITÉ
PAUL VALÉRY
MONTPELLIER 3

ETAPE 3

- Nouvelles données
- Appliquer le modèle

comestibles (C)
non comestibles (N)
toxiques (T)
mortels (M)



<http://www.univ-montp3.fr/miap/ens/info/>

6

UNIVERSITÉ PAUL VALÉRY MONTPELLIER 3

EXEMPLES DE CLASSIFIEURS

Apprentissage 7

UNIVERSITÉ PAUL VALÉRY MONTPELLIER 3

EXEMPLES DE CLASSIFIEURS

- Modèles pour prédire des attributs nominaux ou numériques
- Algorithmes:
 - Arbre de décision
 - Naïve Bayes
 - Support vector machine (SVM)
 - Multi-layer perceptron
 - Bayes network
 - ...
- Meta-classifiers:
 - Combination
 - Bagging
 - Boosting

Apprentissage 8

UNIVERSITÉ PAUL VALÉRY MONTPELLIER 3

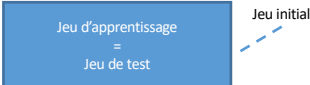
QUEL JEU D'APPRENTISSAGE?

Apprentissage 9

UNIVERSITÉ PAUL VALÉRY MONTPELLIER 3

QUEL JEU D'APPRENTISSAGE?

- Stratégie 1 : Ensemble d'apprentissage = Ensemble de test
 - Non recommandé pour obtenir un **modèle généralisable**
 - Utile pour des tests rapides d'un algorithme : si les performances sont mauvaises lorsque apprentissage = test, pas la peine de continuer...




Apprentissage 10

UNIVERSITÉ PAUL VALÉRY MONTPELLIER 3

QUEL JEU D'APPRENTISSAGE ?

- Stratégie 2 : Ensemble d'apprentissage et de test disjoint
 - Rien ne prouve que la distribution dans le jeu de test soit équivalente à celle du jeu d'apprentissage
 - Utile pour l'apprentissage sur un **flux de données** : définition d'un échantillon pour apprendre et test sur le reste du flux




Apprentissage 11

UNIVERSITÉ PAUL VALÉRY MONTPELLIER 3

QUEL JEU D'APPRENTISSAGE ?

- Stratégie 3 (recommandée) - Validation croisée (Cross-Validation) :
 - La bonne approche pour un **modèle généralisable**
 - Nécessite de définir le nombre de plies k




1. On divise le jeu initial en k échantillons, puis on sélectionne un des k échantillons comme ensemble de test et les (k-1) autres comme ensemble d'apprentissage

Apprentissage 12

UNIVERSITÉ PAUL VALÉRY MONTPELLIER

QUEL JEU D'APPRENTISSAGE ?

- Stratégie 3 (recommandée) - Validation croisée (Cross-Validation) :
- La bonne approche pour un **modèle généralisable**
- Nécessite de définir le nombre de plies k



- On divise le jeu initial en k échantillons, puis on sélectionne un des k échantillons comme ensemble de test et les $(k-1)$ autres comme ensemble d'apprentissage
- On calcule le score de performance

Jeu initial


Apprentissage

13

UNIVERSITÉ PAUL VALÉRY MONTPELLIER

QUEL JEU D'APPRENTISSAGE ?

- Stratégie 3 (recommandée) - Validation croisée (Cross-Validation) :
- Approche pour un **modèle généralisable**
- Nécessite de définir le nombre de plies k



- On divise le jeu initial en k échantillons, puis on sélectionne un des k échantillons comme ensemble de test et les $(k-1)$ autres comme ensemble d'apprentissage
- On calcule le score de performance
- On recommence en sélectionnant un autre échantillon de test parmi les $(k-1)$ échantillons non encore utilisés

Jeu initial


Apprentissage

14

UNIVERSITÉ PAUL VALÉRY MONTPELLIER

QUEL JEU D'APPRENTISSAGE ?

- Stratégie 3 (recommandée) - Validation croisée (Cross-Validation) :
- Approche pour un **modèle généralisable**
- Nécessite de définir le nombre de plies k (ici 3)



- On divise le jeu initial en k échantillons, puis on sélectionne un des k échantillons comme ensemble de test et les $(k-1)$ autres comme ensemble d'apprentissage
- On calcule le score de performance
- On recommence en sélectionnant un autre échantillon de test parmi les $(k-1)$ échantillons non encore utilisés
- L'opération se répète k fois pour que chaque sous-échantillon ait été utilisé exactement une fois comme ensemble de test.
- On moyenne les scores de performance.

Jeu initial


Apprentissage

15

UNIVERSITÉ PAUL VALÉRY MONTPELLIER

QUEL JEU D'APPRENTISSAGE ?

- Stratégie 3 (recommandée) - Validation croisée (Cross-Validation) :
- Approche pour un **modèle généralisable**
- Nécessite de définir le nombre de plies k (ici 3)



- Peut être coûteux en cas de gros volumes de données

Jeu initial


Apprentissage

16

UNIVERSITÉ PAUL VALÉRY MONTPELLIER

QUEL JEU D'APPRENTISSAGE ?

- Stratégie 4 - leave-one-out cross-validation
- Cas particulier : $k=n$



- On apprend sur $(n-1)$ observations
- On teste le modèle sur la i ème observation
- On répète cette opération n fois

Jeu initial

Apprentissage

17

UNIVERSITÉ PAUL VALÉRY MONTPELLIER

PERFORMANCE

Apprentissage

18

MATRICE DE CONFUSION

- Pour une classe donnée, un classifieur et un exemple, on a 4 possibilités :
 - L'exemple est de cette classe et le classifieur ne se trompe pas : **vrai positif**

	Comestible prédite	Comestible non prédite
Exemples Comestibles	Vrais Positifs (VP)	
Exemples non Comestibles		

Apprentissage 19

MATRICE DE CONFUSION

- Pour une classe donnée, un classifieur et un exemple, on a 4 possibilités :
 - L'exemple est de cette classe et le classifieur ne se trompe pas : **vrai positif**
 - L'exemple est de cette classe mais le classifieur se trompe : **faux négatif**

	Comestible prédite	Comestible non prédite
Exemples Comestibles	Vrais Positifs (VP)	Faux Négatifs (FN)
Exemples non Comestibles		

Apprentissage 20

MATRICE DE CONFUSION

- Pour une classe donnée, un classifieur et un exemple, on a 4 possibilités :
 - L'exemple est de cette classe et le classifieur ne se trompe pas : **vrai positif**
 - L'exemple est de cette classe mais le classifieur se trompe : **faux négatif**
 - L'exemple n'est pas de cette classe mais le classifieur la lui attribue : **faux positif**

	Comestible prédite	Comestible non prédite
Exemples Comestibles	Vrais Positifs (VP)	Faux Négatifs (FN)
Exemples non Comestibles	Faux Positifs (FP)	

Apprentissage 21

MATRICE DE CONFUSION

- Pour une classe donnée, un classifieur et un exemple, on a 4 possibilités :
 - L'exemple est de cette classe et le classifieur ne se trompe pas : **vrai positif**
 - L'exemple est de cette classe mais le classifieur se trompe : **faux négatif**
 - L'exemple n'est pas de cette classe mais le classifieur la lui attribue : **faux positif**
 - L'exemple n'est pas de cette classe et le classifieur ne le range pas non plus dans cette classe : **vrai négatif**

	Comestible prédite	Comestible non prédite
Exemples Comestibles	Vrais Positifs (VP)	Faux Négatifs (FN)
Exemples non Comestibles	Faux Positifs (FP)	Vrais Négatifs (VN)

Apprentissage 22

INTERPRETATION DES RESULTATS

- Accuracy : exactitude**
 - $Accuracy = (VP+VN)/(VP+FN+FP+VN)$

Apprentissage 23

INTERPRETATION DES RESULTATS

- Rappel (Recall)**
 - Se calcule à partir de la 1^{ère} ligne de la matrice de confusion
 - $Rappel = VP/(VP+FN)$
 - Rapport entre le nombre d'éléments bien classés et le nombre total d'éléments qui devraient être bien classés
 - Intuition : Si l'on considère tous les champignons comestibles ont été trouvés, est ce que l'on n'en a pas oublié ?
 - Un Rappel de 1 signifie que tous les exemples positifs ont été trouvés

Apprentissage 24

UNIVERSITÉ PAUL VALÉRY MONTPELLIER **INTERPRETATION DES RESULTATS**

- **Précision**
- Se calcule à partir de la 1^{ère} colonne de la matrice de confusion
 - $\text{Précision} = \text{VP} / (\text{VP} + \text{FP})$
- Intuition : *Sur tous les champignons comestibles proposés à l'utilisateur, combien lui en a-t-on proposés qui en fait ne l'intéresseront pas ?*
- Une précision de 1 exprime le fait que tous les exemples classés positifs l'étaient vraiment

Apprentissage 25

UNIVERSITÉ PAUL VALÉRY MONTPELLIER **INTERPRETATION DES RESULTATS**

- **F-Measure**
- Regroupe les deux indicateurs
- Moyenne harmonique du rappel et de la précision
 - $\text{F-Measure} = 2 * \text{Rappel} * \text{Précision} / (\text{Rappel} + \text{Précision})$

Apprentissage 26

UNIVERSITÉ PAUL VALÉRY MONTPELLIER **EXEMPLE : TEST CLINIQUE**

- Etude sur 4000 femmes, 40 ans et plus, apparemment en bonne santé.
- On leur a fait passer 2 tests de dépistage du col de l'utérus :
 - Un examen histologique, plutôt lourd, qui requiert d'être interprété par un expert, et qui servira de vérité terrain.
 - Un frottis de dépistage, qui est un examen beaucoup plus simple et moins invasif, qui sera ici l'analyse de notre algorithme d'apprentissage

	Frottis+	Frottis-
Cancer	190 (VP)	10 (FN)
Pas Cancer	210 (FP)	3590 (VN)

Apprentissage 27

UNIVERSITÉ PAUL VALÉRY MONTPELLIER **EXEMPLE : TEST CLINIQUE**

	Frottis+	Frottis-
Cancer	190 (VP)	10 (FN)
Pas Cancer	210 (FP)	3590 (VN)

- $\text{Exactitude} = (\text{VP} + \text{VN}) / (\text{VP} + \text{FN} + \text{FP} + \text{VN}) = (190 + 3590) / 4000 = 0,945$

Pas mal ?

Apprentissage 28

UNIVERSITÉ PAUL VALÉRY MONTPELLIER **EXEMPLE : TEST CLINIQUE**

	Frottis+	Frottis-
Cancer	190 (VP)	10 (FN)
Pas Cancer	210 (FP)	3590 (VN)

- $\text{Exactitude} = (\text{VP} + \text{VN}) / (\text{VP} + \text{FN} + \text{FP} + \text{VN}) = (190 + 3590) / 4000 = 0,945$
- *Pas mal ?*
- $\text{Rappel} = \text{VP} / (\text{VP} + \text{FN}) = 190 / (190 + 10) = 0,95$
- Frottis : un bon outil de **dépistage**

Apprentissage 29

UNIVERSITÉ PAUL VALÉRY MONTPELLIER **EXEMPLE : TEST CLINIQUE**

	Frottis+	Frottis-
Cancer	190 (VP)	10 (FN)
Pas Cancer	210 (FP)	3590 (VN)

- $\text{Exactitude} = (\text{VP} + \text{VN}) / (\text{VP} + \text{FN} + \text{FP} + \text{VN}) = (190 + 3590) / 4000 = 0,945$
- *Pas mal ?*
- $\text{Rappel} = \text{VP} / (\text{VP} + \text{FN}) = 190 / (190 + 10) = 0,95$
- Frottis : un bon outil de **dépistage**
- $\text{Précision} = \text{VP} / (\text{VP} + \text{FP}) = 190 / (190 + 210) = 0,475$
- Frottis : un mauvais outil de **diagnostic**

Apprentissage 30