



Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

Méthodes d'évaluation

Dr A. DJEFFAL

1^{ère} année Master Informatique de l'Optimisation et de la Décision

2015-2016

www.abdelhamid-djeffal.net



Notion de risque empirique

Risque réel

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

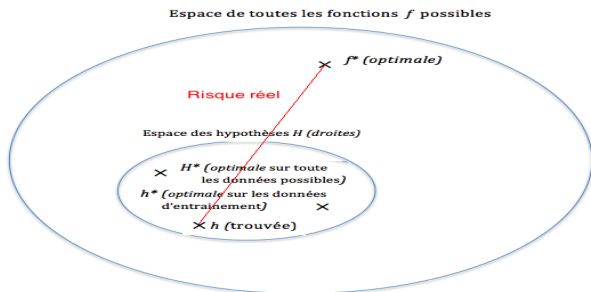
Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Risque réel = $\int_{x,u} L(h(x), f^*(x)) df(x, u)$ (impossible : f inconnue)
- L : fonction de perte





Notion de risque empirique

Risque empirique

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

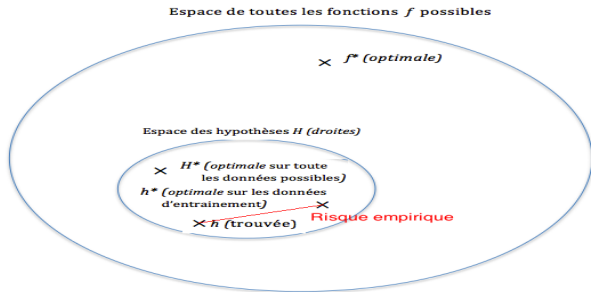
Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Risque empirique = $\frac{1}{n} \sum_{i=1..n} L(h(x_i), u_i)$





Notion de risque empirique

Risque empirique

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

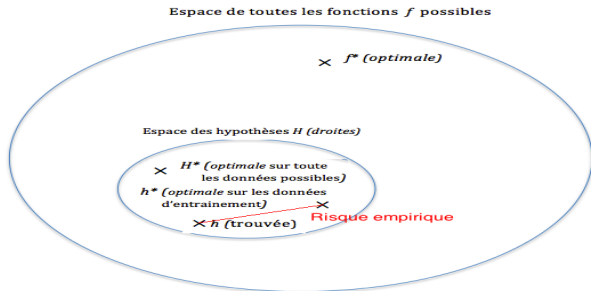
Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Risque empirique = $\frac{1}{n} \sum_{i=1..n} L(h(x_i), u_i)$
- $L(h(x_i), u_i) = 1$ si $h(x_i) \neq u_i$, 0 sinon





Notion de risque empirique

Risque empirique

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

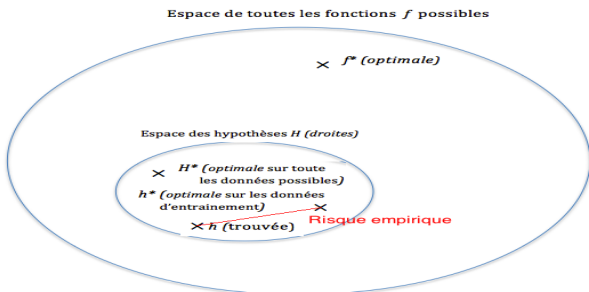
Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Apprentissage : minimisation du risque empirique
- Empirical Risk Minimization (ERM)
- $h = \operatorname{argmin}_H(R_{emp}(H, D))$





Évaluation

Objectif

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Évaluer pour s'assurer de la capacité de généralisation en dehors des données d'entraînement.
- Peut être testé sur :
 - Les données d'entraînement elles-mêmes : Taux élevé mais ne garantit pas une bonne généralisation (sur-apprentissage)
 - Des données différentes : écartés dès le départ de la base d'entraînement



Mesures d'évaluation

Taux de reconnaissance

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Taux de reconnaissance : rapport (pourcentage) entre le nombre de donnée correctement classées et le nombre total des données testées

$$P = \frac{1}{N} \sum_{i=1}^N L(y_i, \hat{f}(x_i))$$

avec :

$$L = \begin{cases} 1 & \text{si } y_i = \hat{f}(x_i) \\ 0 & \text{sinon} \end{cases}$$



Mesures d'évaluation

Matrice de confusion

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Taux d'erreurs = $100 - \textit{précision}$
- Aucune information sur la nature des erreurs,
- Important : quelle classe est considérée comme quelle classe,
- Exemple : considérer un échantillon non cancéreux alors qu'il l'est, est beaucoup plus grave de considérer un échantillon cancéreux alors qu'il ne l'est pas
- Cas de classification binaire :

$$\left\{ \begin{array}{l} \hat{f}(x_i) = +1 \quad \textit{et} \quad y_i = +1 \quad \textit{correcte positive} \\ \hat{f}(x_i) = +1 \quad \textit{et} \quad y_i = -1 \quad \textit{fausse positive} \\ \hat{f}(x_i) = -1 \quad \textit{et} \quad y_i = -1 \quad \textit{correcte négative} \\ \hat{f}(x_i) = -1 \quad \textit{et} \quad y_i = +1 \quad \textit{fausse négative} \end{array} \right.$$



Mesures d'évaluation

Matrice de confusion

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- CP : classe positive considérée positive
- CN : classe négative considérée négative
- FP : classe négative considérée positive
- FN : classe positive considérée négative

Observations (y)	Prédictions (\hat{f})	
	+1	-1
+1	CP	FN
-1	FP	CN



Mesures d'évaluation

Matrice de confusion

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Modèle sans erreurs : $CP + CN = N$
- Multiclasses : nombre de colonnes = nombre de classes
- Précision : $P = \frac{CP+CN}{CP+FP+CN+FN}$



Mesures d'évaluation

Moyenne harmonique

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

Autres mesures :

- Sensitivité : $Sv = \frac{CP}{CP+FN}$
- Spécificité : $Sp = \frac{CN}{CN+FP}$
- Moyenne harmonique = $\frac{2 \times Sv \times Sp}{Sv + Sp}$



Méthodes d'évaluation

Rôle

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Apprentissage \Rightarrow plusieurs paramètres,
 - Exemples d'entraînement,
 - Leur nombre,
 - Paramètres de la méthode d'apprentissage,
 - ...
- Choix des paramètres \Rightarrow plusieurs essais
- Meilleurs paramètres : Précision = 100 %
- L'évaluation permet de tirer des conclusions sur le comportement d'un modèle face à tout l'espace d'exemples



Méthodes d'évaluation

Méthode HoldOut

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Suppose que les données d'entraînement sont tout l'espace d'exemples
- Diviser l'ensemble des données en deux parties :
 - ① Partie d'entraînement
 - ② Partie de test
- Tester l'efficacité du modèle ne dehors des données d'entraînement
- Question : quels pourcentages
- Le pire : positifs dans une partie et négatifs dans l'autre



Méthodes d'évaluation

Validation croisée

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Minimiser l'influence du choix du partitionnement dans HoldOut
- Principe : Subdiviser D l'ensemble d'apprentissage en k sous ensemble disjoints D_1, D_2, \dots, D_k de même taille
- A l'itération i le sous-ensemble D_i est réservé pour le test et le reste des exemples sont utilisés pour l'entraînement le modèle.
- Précision = moyenne des k précisions
- Leave-One-Out : cas particulier $k = n$



Méthodes d'évaluation

Le Bootstrap

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Bootstrap ou échantillonnage par remplacement
- Choisir N exemples aléatoirement de D
- Puis entraîner le modèle sur ces N exemples
- des exemples peuvent être choisis plus d'une fois, et d'autres pas du tout
- Tester sur Les exemples non choisis
- Répéter k fois pour obtenir une précision moyenne
- Précision = $P = \sum_{i=1}^k (0.632 \times P_{i_{test}} + 0.368 \times P_{i_{entr}})$
- $P_{i_{test}}$: précision du modèle dans l'itération i sur les données de test
- $P_{i_{entr}}$: précision du modèle dans l'itération i sur les données d'entraînement



Combinaison de modèles

Combinaison

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

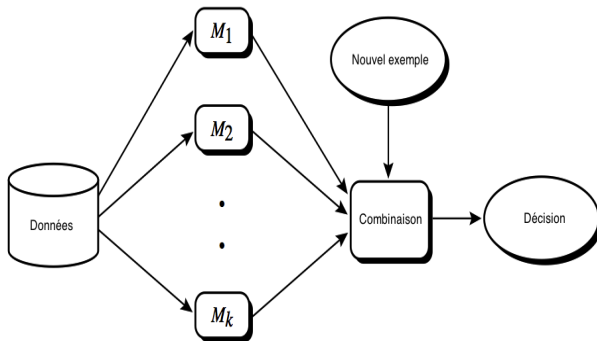
Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles





Combinaison de modèles

Bagging

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- Subdivise l'ensemble D en n sous-ensembles.
- À partir de chaque sous-ensemble D_i , on apprend un modèle M_i
- L'ensemble de ces modèles forme un modèle composé M_* .
- Pour classer un nouvel exemple :
 - Il est exposé à chaque modèle M_i pour obtenir une classe c_{M_i} .
 - Chaque décision est considérée comme un vote.
 - La classe de décision est prise comme la classe la plus votée.



Combinaison de modèles

Boosting

Méthodes
d'évaluation

Dr A.
DJEFFAL

Notion de
risque
empirique

Estimation du
risque

Évaluation

Mesure
d'évaluation

Méthodes
d'évaluation

Combinaison
de modèles

- On associe des poids aux exemples.
- Une série de k modèles est itérativement apprise.
- Après qu'un modèle M_i est construit, les poids des exemples sont mis à jour de telle sorte à attirer l'attention du modèle M_{i+1} aux exemples mal-classés par le modèle M_i .
- Le Modèle final M_* combine les voltes des k modèles pondérés par leur précisions.