

— Classification supervisée —

Syllabus du cours

Romain Azaïs

0 – CONTEXTE

- 0.1 Vision statistique
- 0.2 Vision apprentissage
- 0.3 Et si Y n'est pas binaire?

One-vs-One, One-vs-All

1 – RISQUE, RISQUE BAYÉSIEN

- 1.1 Modèle probabiliste
- 1.2 Risque d'un classifieur
- 1.3 Exemple : mélange de gaussiennes
- 1.4 Estimation des paramètres : approche plug-in

LDA, QDA
lien avec k-means

2 – D'AUTRES CLASSIFIEURS À FRONTIÈRE LINÉAIRE

- 2.1 Régression logistique
- 2.2 Perceptron
- 2.3 Machines à vecteurs supports
 - 2.3.1 Classes linéairement séparables
 - 2.3.2 Introduction d'erreurs
 - 2.3.3 Astuce du noyau
 - 2.3.4 Digression : données structurées

régression logistique
perceptron
SVM
hard margin
soft margin

3 – K PLUS PROCHES VOISINS

- 3.1 Définition
- 3.2 Un résultat pour éliminer la statistique d'ordre
- 3.3 Approximation du classifieur des kNN
- 3.4 Limite du risque moyen
- 3.5 Comportement lorsque $k \rightarrow \infty$

kNN

4 – BOOSTING

- 4.1 Apprentissage à partir de données pondérées
- 4.2 Algorithme AdaBoost
- 4.3 Risque empirique

AdaBoost

5 – MINIMISATION DU RISQUE EMPIRIQUE : APPRENTISSAGE PAC

- 5.1 L'idée
- 5.2 Cas de la séparation déterministe
- 5.3 Borne agnostique
- 5.4 Application : MVC

famille finie de classifieurs

MVC

Références

The Elements of Statistical Learning. J.H. Friedman, R. Tibshirani et T. Hastie (2001).

Lectures on the Nearest Neighbor Method. G. Biau et L. Devroye (2015).

Statistique et apprentissage : notes de cours à l'X de G. Fort, M. Lerasle et É. Moulines (2020).

Fondamentaux de l'apprentissage statistique : notes de cours à Orsay de S. Arlot (2020).

Machine learning : slides on Boosting à Carnegie Mellon de T. Mitchell et Z. Bar-Joseph (2012).

Apprentissage statistique : slides sur les SVM à Montpellier de J. Salmon et N. Verzelen (2020).