

Introduction à l'apprentissage supervisé

ECTS : 3

Volume horaire : 21

Description du contenu de l'enseignement :

- 1- Analyse factorielle discriminante
- 2- Analyse discriminante linéaire et quadratique
- 3- Classification bayésienne à l'aide de modèles de mélange
- 4- Classifieur bayésien et classifieur bayésien naïf
- 5- Sélection de modèles de mélange parcimonieux
- 6- Arbres de décision
- 7- Forêts aléatoires

L'ensemble de ces méthodes enseignées est illustré par des démonstrations du logiciel R sur des jeux de données réel (principalement Analyse Discriminante linéaire et quadratique, Classification bayésienne gaussienne, Classifieur bayésien naïf, Forêts aléatoires).

Compétence à acquérir :

Ce cours présente les méthodes élémentaires d'apprentissage supervisé suivantes : analyse factorielle discriminante, classification bayésienne à l'aide de modèles de mélange, arbres de décision et forêts aléatoires. Les propriétés théoriques et différentes formulations de ces méthodes sont présentées. Leurs mises en oeuvre, ainsi que celles de leurs variantes, sont illustrées à l'aide de traitements de données effectués avec le logiciel R. L'objectif de ce cours est l'acquisition de la maîtrise de ces méthodes élémentaires d'apprentissage supervisé.

Mode de contrôle des connaissances :

Examen

Bibliographie, lectures recommandées :

- Benzecri, J.-P. (1980) Pratique de l'analyse des données. Dunod. Paris.
- Bouveyron, C., Celeux, G., Murphy, T., & Raftery, A. (2019) Model-Based Clustering and Classification for Data Science: With Applications in R, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge: Cambridge University Press.
- Breiman,L., Friedman, J.H., Olshen,R., and Stone, C.J. (1984). Classification and Regression Trees, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific California.
- Hastie, T., Tibshirani, R., Friedman, J. (2009) The Elements of Statistical Learning : Data Mining, Inference, and Prediction, Second Edition, Springer Series in Statistics.
- James, G., Witten, D., Hastie, T., Tibshirani, R., Taylor, J., (2023) An Introduction to Statistical Learning: With Applications in Python, Springer International Publishing.
- Lebart, L., Piron, M. , Morineau, A. (2006) (4ème édition, refondue) Statistique Exploratoire Multidimensionnelle, 480 pages, Dunod.
- Saporta, G. (2006), Probabilités, Analyse des données et Statistique, 656 pages, Technip.