STA217 - Apprentissage statistique en grande dimension

Mis à jour le 19-02-2026

Code : STA217

Unité d'enseignement de type cours

9 crédits

Volume horaire de référence (+/- 10%) : **70 heures**

Responsabilité nationale : EPN06 - Mathématique et statistique / N'deye NIANG

Contact national :

EPN06 Mathématiques et statistiques

2 rue conté
Accès 35 3 ème étage porte 19
75003 Paris

Anne - Solenne Maroulle
anne-solenne.marroulle@lecnam.net

Présentation

Prérequis

Avoir le niveau des unités d'enseignement : STA201, STA211. Avoir de bonnes connaissances en calcul matriciel, probabilités et inférence statistique.

Objectifs pédagogiques

Maîtriser les méthodes d'apprentissage statistique pour l'analyse de données de grande dimension

Compétences

Maîtriser les méthodes d'apprentissage statistique pour l'analyse de données structurées ou non mettant en jeu un grand nombre de variables et/ou d'individus

Programme

Contenu

1. Définition et enjeux de l'apprentissage statistique
2. Présentation des problématiques liées à la grande dimension
3. Méthodes statistiques pour la grande dimension
    1. **Méthode non supervisées**
        1. Analyse factorielle sparse
        2. Méthode de classification : Sparse Kmeans, Subspace clustering, Consensus de partitions
        3. Méthodes descriptives pour données multiblocs
        4. ACP fonctionnelle
        5. Classification de données fonctionnelles
    2. **Méthodes supervisées**
        1. Méthodes de sélection de variables en grande dimension
        2. Méthodes de régularisation
        3. Méthodes de classification supervisée
        4. Méthodes supervisées pour données multiblocs
        5. Méthodes de régression pour données fonctionnelles
    3. **Méthodes Clusterwise** pour un grand nombre d'observations hétérogènes

Modalités de validation

- Projet(s)
- Examen final

Description des modalités de validation

Un examen écrit + un projet personnel sanctionneront la fin des cours.
Le projet personnel devra mettre en application les techniques décrites en cours. Il pourra faire l'objet d'un présentation orale

Bibliographie

Titre	Auteur(s)
Functional Data Analysis. Second Edition (Springer Series in Statistics, 2005)	Ramsay J.O. and Silverman B.W.
Applied Functional Data Analysis (Springer, 2002)	Ramsay J.O. and Silverman B.W.

Functional Data Analysis with R. (Chapman & Hall, 2024)	Crainiceanu C.M., Goldsmith J., Leroux A. and Cui E.
Analyse des données (Hermes,2003)	G.GOVAERT
Statistique explicative appliquée (Technip, 2003)	J.P.NAKACHE, J.CONFAIS
Probabilités, analyse des données, statistique 3ème édition (Technip,2011))	G.SAPORTA
The Elements of Statistical Learning (Springer 2009). Téléchargeable ici : https://www.sas.upenn.edu/~fdiebold/NoHesitations/BookAdvanced.pdf	T.HASTIE, J.FRIEDMAN, F.TIBSHIRANI
An introduction to statistical learning: With applications in python. (Springer Nature 2023). Téléchargeable ici:https://www.statlearning.com/	James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J.
An introduction to statistical learning: With applications in R. (Springer Nature 2021). Téléchargeable ici:https://www.statlearning.com/	James, G., Witten, D., Hastie, T., & Tibshirani, R.
Mining of massive data sets. (Cambridge university press 2020)	Leskovec, J., Rajaraman, A., & Ullman, J. D.
Multivariate Analysis, (Academic Press, 1979)	K.V. Mardia, J.T. Kent and J.M. Bibby