

Analyse de données : Apprentissages non supervisé et supervisé. Simulation stochastique.



École / Prépa
ENSEIRB-
MATMECA

Présentation

Code interne : EMM7-ALEA1

Description

Dans ce cours, nous aborderons différentes techniques d'apprentissage statistique. Plus précisément nous nous intéresserons à l'apprentissage non supervisé avec les méthodes d'analyse en composantes principales et de partitionnement et à l'apprentissage supervisé avec les méthodes de régression et de classification. Ces méthodes seront mises en oeuvre avec le langage de programmation python. Ensuite, des mini projets seront proposés.

Heures d'enseignement

CM	Cours Magistraux	2h
TDM	Travaux Dirigés sur Machine	16h
CI	Cours Intégrés	12h

Syllabus

I- Introduction à l'analyse de données

II- Apprentissage non supervisé : 1. Analyse en composantes principales (ACP) 2. Partitionnement/clustering de données (dendrogrammes, k-means)

III- Apprentissage supervisé : 1. Classification (plus proches voisins, forêts aléatoires, Bayésien, machines à vecteurs de supports (SVM)) 2. Régression (linéaire, maximum de vraisemblance).

IV- Notion générales d'apprentissage statistique : validation, surapprentissage, généralisation

Informations complémentaires

ENSEIRB-MATMECA

Apprentissage non supervisé : ACP , partitionnement. Apprentissage supervisé : régression , classification. Filtrage, classification, estimateurs, apprentissage.

Modèles et simulation stochastique: méthodes de Monte-Carlo, algorithmes stochastiques

Langage R. Matlab

Modalités de contrôle des connaissances

Évaluation initiale / Session principale

Type d'évaluation	Nature de l'évaluation	Durée (en minutes)	Nombre d'épreuves	Coefficient de l'évaluation	Note éliminatoire de l'évaluation	Remarques
Projet	Soutenance			1		

Infos pratiques

Contacts

Nicolas Papadakis

✉ Nicolas.Papadakis@bordeaux-inp.fr