

Analyse de données.



Présentation

Code interne : EI9IF344

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 lors de 2 séances de TP avec le langage de programmation R. Ensuite, des mini projets vous seront proposés.

Syllabus

Introduction à l'analyse des données et à l'apprentissage statistique

Apprentissage non supervisé :

- * Analyse en Composantes Principales (ACP)
- * Clustering (kmeans et classification ascendante hiérarchique)

Apprentissage supervisé :

- * Régression linéaire simple et multiple
- * classification (knn, analyse discriminante linéaire et quadratique, bayésien naïf, régression logistique, Arbres de décision et forêts aléatoires).

Mise en œuvre des méthodes et méthodologies avec le logiciel R

Informations complémentaires

Apprentissage non supervisé : ACP , partitionnement. Apprentissage supervisé : régression , classification.

Filtrage, classification, estimateurs, apprentissage. Langage R.

Modalités de contrôle des connaissances



Évaluation initiale / Session principale - Épreuves

Type d'évaluation	Nature de l'épreuve	Durée (en minutes)	Nombre d'épreuves	Coefficient de l'épreuve	Note éliminatoire de l'épreuve	Remarques
Contrôle Continu	Participation Active			1		
Contrôle Continu	Compte-Rendu			1		
Projet	Rapport			1		
Projet	Soutenance			1		

Infos pratiques

Contacts

Marie Chavent

✉ Marie.Chavent@bordeaux-inp.fr