STA218 - Science de la donnée en milieu professionnel

Mis à jour le 18-02-2026

Code : STA218

Unité d'enseignement de type mixte

3 crédits

Volume horaire de référence (+/- 10%) : **30 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

Bonne connaissance des méthodes avancées de fouille de données, connaissance d'au moins un langage de programmation. Niveau recommandé UE STA201, STA211, RCP216

Objectifs pédagogiques

Maîtriser le processus d'extraction de connaissances à travers des techniques avancées d'analyse des données autour de problématiques professionnelles dans différents domaines d'application : biologie, environnement, banque, finance, assurance, marketing, etc.

Programme

Contenu

- Identification des problématiques liées aux données de l'étude
- Études préalables et mise en place du processus de fouille
  - Prétraitement des données : Analyse de la qualité des données, gestion des valeurs manquantes ou aberrantes
  - identification des techniques de fouille adaptées
- Mise en œuvre des méthodes de fouille en intégrant les contraintes métiers (biais, coût d'acquisition des données, etc)

Modalités de validation

- Contrôle continu
- Projet(s)

Description des modalités de validation

La validation de l'UE se fera sur la base d'un projet qui pourra donner lieu à une soutenance orale.

Bibliographie

Titre	Auteur(s)
An Introduction to Statistical Learning (2021) with applications in R Téléchargeable ici: https://www.statlearning.com/	James, Witten, Hastie, & Tibshirani
An Introduction to Statistical Learning with applications in Python (2023) Téléchargeable ici: https://www.statlearning.com/	James, Witten, Hastie, Tibshirani & Taylor
Etude de cas en statistique décisionnelle (Technip, 2009)	S.TUFFERY