

Fouille de données et apprentissage automatique 1



Niveau d'étude
BAC +4



ECTS
3 crédits



Composante
École
d'ingénieur
Denis Diderot



Période de
l'année
Semestre 2

En bref

- › **Langue(s) d'enseignement:** Français
- › **Méthode d'enseignement:** En présence
- › **Forme d'enseignement :** Cours magistral, Travaux dirigés & Travaux pratiques
- › **Ouvert aux étudiants en échange:** Oui

Présentation

OBJECTIFS

Savoir utiliser les bibliothèques python scikit-learn, numpy, pandas. Savoir concevoir un modèle prédictif, évaluer sa performance, interpréter les résultats sur le plan biologique. Savoir restituer les méthodes et résultats et répondre à des questions méthodologiques et biologiques.

SYLLABUS

Maîtriser les concepts et méthodes d'apprentissage automatique supervisé. Comprendre les fondamentaux et les limitations de chaque approche.

- Les notions fondamentale d'apprentissage automatique supervisé : inférence inductive, fonction de perte, évaluation de la performance du modèle, sous- et sur-apprentissage, complexité du modèle ;

- Les modèles prédictifs basiques : régression logistique, classifieur bayésien naïf, K plus proches voisins ;
- Algorithmes de sélection de modèles par pénalisation : ridge, lasso, elastic net ;
- Les arbres de décision et leurs ensembles : random forest, bagging, boosting ;
- Machines à support vectoriel ;
- Projet de développement d'un outil de prédiction pour les données biomédicales.

En bref

LIEU(X)

- › Campus des Grands Moulins

Pour en savoir plus, rendez-vous sur > u-paris.fr/choisir-sa-formation