

Exploration de données biomédicales à grande échelle 1



Niveau d'étude
BAC +3 (niveau 6)



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 effectuer une analyse exploratoire des données et produire une visualisation informative aussi bien qu'interpréter ces résultats.

SYLLABUS

Comprendre les spécificités des données en grande dimension. Apprendre les bases des méthodes utilisées pour leur exploration et visualisation à l'aide d'apprentissage automatique :

- Introduction des concepts généraux d'analyse de données, fouille de données et d'apprentissage automatique non-supervisé

- Analyse exploratoire des données avec python. Pandas pour détection des point aberrantes, analyses de corrélations, analyse exploratoire uni- et multi-varié (suite de biostatistique)

- Méthodes de réduction de dimension linéaires et non-linéaires: Analyse en Composantes Principales, Multidimensional Scaling, Isomap, local linear embedding, Analyse en Composantes indépendantes, modèles à facteurs latents.

- Méthodes avancées de visualisation des données à grande échelle : tSNE, UMAP, Diffusion Maps. Leur principe, exemples d'utilisation et limitations.

- Réalisation d'un projet d'analyse de données en grande dimension.

En bref

LIEU(X)

- › Campus des Grands Moulins

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