

ENG-209

Data science pour ingénieurs avec Python

Bouillet Eric Pierre, Pellet Jean-Philippe, Verscheure Olivier

Cursus	Sem.	Type
Génie électrique et électronique	BA3	Obl.
HES - EL	H	Opt.

Langue d'enseignement	français
Crédits	3
Session	Hiver
Semestre	Automne
Examen	Pendant le semestre
Charge	90h
Semaines	14
Heures	3 hebdo
Cours	1 hebdo
Projet	2 hebdo
Nombre de places	

Résumé

Ce cours propose une immersion progressive et complète dans le domaine de la Data Science à travers le langage Python. Il guide les étudiants depuis la manipulation de données brutes jusqu'à la modélisation et l'extraction de connaissances utiles à partir de données réelles.

Contenu**1. Data Engineering : Manipulation et visualisation des données**

- Introduction aux spécificités du langage Python
- Prise en main des environnements de développement (IDE) dédiés à la Data Science
- Utilisation des bibliothèques Python pour la manipulation (ex. pandas, NumPy) et la visualisation (ex. Matplotlib, Seaborn) des données (avec exercices)

2. Évaluation intermédiaire (1/3)**3. Data Science : Modélisation**

- Introduction à la modélisation et présentation des principales catégories de modèles
- Régression linéaire et polynomiale (avec exercices)
- Classification supervisée : régression logistique et arbres de décision (avec exercices)
- Techniques de validation de modèles - ex. cross-validation, métriques d'évaluation (avec exercices)

4. Projet final (2/3)

- Un projet appliqué mobilisant l'ensemble des compétences acquises pour résoudre un problème réel de Data Science.

Mots-clés

Programmation, Python, Data Science, Machine Learning

Compétences requises**Concepts importants à maîtriser**

Bases de programmation

- Variables, types de données, boucles, conditions, fonctions
- Structures de données : listes, dictionnaires, tableau

Notions fondamentales en mathématiques

- Algèbre de base (résolution d'équations simples, manipulation de formules)
- Notions de fonctions, dérivées et courbes
- Compréhension des vecteurs et matrices

Statistiques descriptives de base et éléments de probabilité

Acquis de formation

A la fin de ce cours l'étudiant.e doit être capable de:

- Modéliser en langage Python les éléments pertinents d'une situation
- Transposer les connaissances acquises en C++ vers Python
- Appliquer les constructions de haut niveau de Python pour créer des structures de données linéaires ou associatives
- Utiliser les bibliothèques Python courantes pour manipuler et visualiser des données
- Appliquer des méthodes de modélisation supervisée à des jeux de données réels
- Créer une analyse de données complète dans le cadre d'un projet structuré

Compétences transversales

- Utiliser une méthodologie de travail appropriée, organiser un/son travail.
- Accéder aux sources d'informations appropriées et les évaluer.
- Utiliser les outils informatiques courants ainsi que ceux spécifiques à leur discipline.
- Faire preuve d'esprit critique

Méthode d'enseignement

Séances ex cathedra avec exercices et travaux pratiques

Travail attendu

Participation aux cours, résolution d'exercices, réalisation du travail attendu en dehors de heures de cours

Méthode d'évaluation

Contrôle continu pendant le semestre réparti selon:

- Evaluation intermédiaire: 1/3 de la note totale
- Projet final avec rapport: 2/3 de la note totale

Encadrement

Office hours	Non
Assistant.e.s	Oui
Forum électronique	Oui

Ressources

Service de cours virtuels (VDI)

Oui

Liens Moodle

- <https://go.epfl.ch/ENG-209>