

Utilisation & Exploitation des Données Introduction to Data Science			
Code ECUE Course code: UED		UE (Crédits ECTS de l'UE) : UE2-4 (8 ECTS)	
Département Department	IA	Cours Lectures	8h45
Coordonnateurs Lecturers	B. Chardin	T.D. Tutorials	12h30
Période Year of study	A1	T.P. Laboratory sessions	15h00
Semestre Semester	S2	Projet Project	
Evaluation Assessment method(s)	1 écrit, Travaux Pratiques	Non encadré Unsupervised	
Langue d'instruction Language of instruction	Français	Horaire global Total hours	36h15
Type de cours Type of course	Obligatoire	Travail personnel Homework	13h00
Niveau Level of course	Premier cycle universitaire		

Compétences attendues :

- Extraire de l'information de données structurées et semi-structurées.
- Automatiser et systématiser les traitements d'analyse des données.
- Connaître les principes de fonctionnement de l'apprentissage automatique
- Savoir mettre en œuvre et évaluer une méthode d'apprentissage supervisé

Pré-requis : Information Numérique (S1)

Contenu :

- Structures de données (collections)
- Transformation et préparation des données
- Extraction d'information et analyse
- Données structurées (CSV, relations) et semi-structurées (HTML, XML, JSON)
- Implémentation en Python (pandas, BeautifulSoup) et en SQL
- Apprentissage supervisé et techniques usuelles associées
- kNN, SVM, SVR, arbres et forêts de décision
- Méthodologie d'évaluation et de validation

Bibliographie :

- Wes McKinney. Python for Data Analysis. Second edition. O'Reilly Media, Inc. ISBN: 9781491957660. 2017.
- Philipp K. Janert. Data Analysis with Open Source Tools. O'Reilly Media, Inc. ISBN: 9780596802356. 2010.
- R. Lyman Ott and Michael Longnecker. An Introduction to Statistical Methods and Data Analysis. Fifth edition. Duxbury. ISBN: 0534251226. 2001.
- Géron, A. (2019). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. ISBN: 9781492032649.
- Müller, A. C., & Guido, S. (2016). Introduction to machine learning with Python: a guide for data scientists. ISBN: 9781449369415.
- Molnar, C. (2022). Interpretable Machine Learning. ISBN: 9798411463330.
- Patel, A. (2019). Hands-On Unsupervised Learning Using Python. ISBN: 9781492035640.

Expected competencies:

- Discover knowledge from structured and semi-structured data.

- Design and implement data processing utilities.
- Understand the underlying mechanisms of machine learning
- Apply supervised learning solutions from the state of the art and evaluate their performance

Prerequisites: Digital information (S1)

Content:

- data structures (collections)
- data wrangling
- knowledge discovery and data analysis
- structured (CSV, relational data) and semi-structured (HTML, XML, JSON) data
- Python (pandas, BeautifulSoup) and SQL implementations
- Supervised machine learning and standard algorithms
- kNN, SVM, SVR, decision trees and random forests
- Performance evaluation methodology
- Reinforcement learning

Recommended reading:

- Wes McKinney. Python for Data Analysis. Second edition. O'Reilly Media, Inc. ISBN: 9781491957660. 2017.
- Philipp K. Janert. Data Analysis with Open Source Tools. O'Reilly Media, Inc. ISBN: 9780596802356. 2010.
- R. Lyman Ott and Michael Longnecker. An Introduction to Statistical Methods and Data Analysis. Fifth edition. Duxbury. ISBN: 0534251226. 2001.
- Géron, A. (2019). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. ISBN: 9781492032649.
- Müller, A. C., & Guido, S. (2016). Introduction to machine learning with Python: a guide for data scientists. ISBN: 9781449369415.
- Molnar, C. (2022). Interpretable Machine Learning. ISBN: 9798411463330.
- Patel, A. (2019). Hands-On Unsupervised Learning Using Python. ISBN: 9781492035640.