

## Machine Learning, big data and quantitative spatial analysis

**Niveau : 2<sup>ème</sup> année de Master – Master 2**

**Cursus : Data Science et Management de l'Innovation – Political Engineering**

**Semestre : 1<sup>er</sup> semestre**

**Volume horaire : 12 heures de CM, incluant des temps de pratique**

**Intervenant : Jean-Philippe PEYRACHE**

**Contact : [jean.philippe.peyrache@univ-st-etienne.fr](mailto:jean.philippe.peyrache@univ-st-etienne.fr)**

### Objectif du cours :

Les objectifs du cours sont de donner aux étudiant-e-s les éléments de compréhension indispensables au domaine du Machine Learning, ainsi que de les familiariser avec différentes familles d'algorithmes. Le principe général et les différents cas d'usage permettront de distinguer les approches qui peuvent être utilisées en fonction des problématiques rencontrées et le cours sera ensuite centré sur l'apprentissage supervisé dans les tâches de classification, afin de développer le cadre théorique et les difficultés qui peuvent limiter le succès desdites approches.

### Plan de cours :

Introduction au Machine Learning : principe, cas d'usage et différentes approches

La question des données : différences de distribution, difficultés et dimensionnalité

L'apprentissage supervisé : cadre théorique et algorithmes de base

Arbres de décision et forêts aléatoires (méthodes ensemblistes)

Modèles linéaires et introduction à l'apprentissage profond

### Compétences développées :

Les compétences développées par les étudiant-e-s à l'issue du cours sont les suivantes :

- Compréhension générale des problématiques de Machine Learning, de leur utilité et de leurs limites
- Connaissance du cadre théorique lié à l'apprentissage supervisé
- Connaissance des forces et faiblesses de différents algorithmes de classification dans un cadre d'apprentissage supervisé, capacité à les mettre en œuvre sur des jeux de données

### Modalité d'évaluation :

Compte-rendu d'une mise en œuvre pratique de plusieurs algorithmes sur des problèmes déterminés

### Prérequis (à mentionner uniquement s'il y en a) :

Connaissance de base de langage Python et environnement de travail fonctionnant avec Python et la bibliothèque Scikit-learn

### Bibliographie / référence :

*Sur la théorie de l'apprentissage*

Vladimir Vapnik. *Statistical learning theory*. Wiley, 1998.

- Leslie G. Valiant. *A Theory of the Learnable*. *Commun. ACM*, vol. 27, no. 11, pages 1134–1142, 1984.

- Antoine Cornuéjols & Laurent Miclet. *Apprentissage artificiel : concepts et algorithmes*. 2010

*Sur la bibliothèque Python Scikit-Learn :*

- <https://scikit-learn.org/stable/index.html>