

# Python, machine learning

3 j (21 heures)

Ref : MCEP

## Public

Développeur, data engineer, data analyst, data scientist, chercheur, ingénieur R&D, chef de projet technique, statisticien et toute personne travaillant dans la data et sachant manipuler du code informatique

## Pré-requis

Connaitre un langage de programmation, idéalement python

## Moyens pédagogiques

Formation réalisée en présentiel ou à distance selon la formule retenue  
Exposés, cas pratiques, synthèse, assistance post-formation pendant trois mois  
Un poste par stagiaire, vidéoprojecteur, support de cours fourni à chaque stagiaire

## Modalités de suivi et d'évaluation

Feuille de présence émargée par demi-journée par les stagiaires et le formateur  
Exercices de mise en pratique ou quiz de connaissances tout au long de la formation permettant de mesurer la progression des stagiaires  
Questionnaire d'évaluation de la satisfaction en fin de stage  
Auto-évaluation des acquis de la formation par les stagiaires  
Attestation de fin de formation

Dans cette formation vous comprendrez les algorithmes les plus efficaces de Machine Learning. Vous entraînerez vos premiers modèles prédictifs sur des jeux de données issus de cas réels. Vous apprendrez à les évaluer, à les optimiser, et à éviter les écueils classiques de ce type de projet.

A travers de nombreuses mises en pratique et illustrations, vous comprendrez exactement les cas d'usage à forte valeur ajoutée qui peuvent être adressés par ces technologies, et comment les mettre en œuvre en Python avec la librairie scikit-learn.

## Objectifs

- Comprendre les enjeux et les fondamentaux du Machine Learning
- Apprendre à exploiter les algorithmes de Machine Learning
- Construire des modèles prédictifs à partir de vos données

## Programme détaillé

### INTRODUCTION

---

Python, machine learning

Définition du Machine Learning  
Principaux cas d'usage  
Classification et régression  
Machine Learning supervisé et non supervisé

## **PYTHON ET LE MACHINE LEARNING**

---

Librairies scientifiques : numpy, pandas, et matplotlib  
Scikit-learn  
Serveur Jupyter et notebooks

## **ALGORITHMES DE MACHINE LEARNING SUPERVISES**

---

Régression linéaire et régression logistique  
K plus proches voisins : KNN  
Arbres de décision  
Forêts aléatoires

## **QUANTIFIER LA QUALITE D'UN MODELE**

---

Méthode pour mesurer la qualité  
Critères d'évaluation pour la régression : MAE, MSE, R<sup>2</sup>...  
Critères d'évaluation pour la classification : accuracy, F1 score, ROC AUC...

## **AMELIORER LA QUALITE DES MODELES**

---

Tester plusieurs modèles  
Optimiser le meilleur modèle  
Optimiser les prédictions

## **MENER UN PROJET DE MACHINE LEARNING**

---

Etapes clés du projet  
Packager et déployer le modèle  
Eviter les erreurs classiques  
Superviser le modèle en production

## **ALGORITHMES DE MACHINE LEARNING NON SUPERVISES**

---

K-Means  
DBScan  
Isolation Forests

---