

Durée : 3 jours soit 21 heures

Référence : IF-PyML1

**Public visé :**

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 :**

Pour suivre cette formation, les apprenants doivent connaître un langage de programmation, idéalement python.

**Objectifs pédagogiques :**

Dans cette formation vous apprendrez à :

- Analyser et modéliser des jeux de données
- Entraîner un modèle prédictif et à le déployer.

L'alternance de théorie et de pratique permet de :

- Comprendre les algorithmes utilisés
- Constaté leurs principales propriétés sur des jeux de données réels.

A l'issue de la formation vous saurez :

- Identifier les cas d'usage exploitables par le Machine Learning,
- Comprendre le processus projet à mettre en œuvre
- Créer et utiliser vos premiers modèles.

**Compétences acquises à l'issue de la formation :**

- Analyser et modéliser des jeux de données
- Entraîner un modèle prédictif et à le déployer.
- Comprendre les algorithmes utilisés
- Constaté leurs principales propriétés sur des jeux de données réels.
- Identifier les cas d'usage exploitables par le Machine Learning
- Comprendre le processus projet à mettre en œuvre
- Créer et utiliser vos premiers modèles

**Modalités pédagogiques :**

Session dispensée en présentiel ou téléprésentiel, selon la modalité inter-entreprises ou intra-entreprises sur mesure.

La formation est animée par un(e) formateur(trice) durant toute la durée de la session et présentant une suite de modules théoriques clôturés par des ateliers pratiques validant l'acquisition des connaissances. Les ateliers peuvent être accompagnés de Quizz.

L'animateur(trice) présente la partie théorique à l'aide de support de présentation, d'animation réalisée sur un environnement de démonstration.

En présentiel comme en téléprésentiel, l'animateur(trice) accompagne les participants durant la réalisation des ateliers.

**Moyens et supports pédagogiques :**

**Cadre présentiel**

Salles de formation équipées et accessibles aux personnes à mobilité réduite.

- Un poste de travail par participant
- Un support de cours numérique ou papier (au choix)
- Un bloc-notes + stylo
- Vidéoprojection sur tableau blanc
- Connexion Internet
- Accès extranet pour partage de documents et émargement électronique

**Cadre téléprésentiel**

Session dispensée via notre solution iClassroom s'appuyant sur Microsoft Teams.

- Un compte Office 365 par participant
- Un poste virtuel par participant
- Un support numérique (PDF ou Web)
- Accès extranet pour partage de documents et émargement électronique

**Informations sur l'accessibilité :**



## Description / Contenu

### Module 1 : Qu'est-ce que le Machine Learning ?

- Introduction au Machine Learning
- Exemples d'utilisation dans différents secteurs
- Identification des problématiques adressées par le Machine Learning
- Les différents types de Machine Learning : régression, classification, supervisé, non supervisé

### Module 2 : Les librairies Python scientifique :

- Les basiques : Numpy, pandas, et matplotlib
- La librairie Scikit-learn
- Les notebooks Jupyter

### Module 3 : Algorithmes de Machine Learning supervisés, théorie et pratique :

- Régression linéaire polynomiale
- Régression logistique
- K plus proches voisins (KNN)
- Arbres de décision
- Random Forests

### Module 4 : Evaluer un modèle prédictif :

- Comment mesurer la qualité d'un modèle
- Principaux critères d'évaluation
- Créer son propre critère d'évaluation

### Module 5 : Optimiser ses modèles prédictifs :

- Grille d'optimisation
- Recherche aléatoire
- Optimisation de la valeur seuil
- Eviter l'overfitting

### Module 6 : Processus projet :

- Les étapes d'un projet de Machine Learning
- Déployer le modèle en production
- Superviser et mettre à jour le modèle

### Module 7 : Algorithmes de Machine Learning non supervisés :

- Clustering : K-means et DBScan
- Détection d'anomalie : Isolation Forests