

## Formation Machine learning : implémentation en Python

■ <b>Durée :</b>	5 jours (35 heures)
■ <b>Tarifs inter-entreprise :</b>	3 475,00 € HT (standard) 2 780,00 € HT (remisé)
■ <b>Public :</b>	Développeurs Python, Data Scientists, Managers de projets
■ <b>Pré-requis :</b>	Maîtrise des bases du langage Python
■ <b>Objectifs :</b>	Comprendre et choisir une méthode d'apprentissage automatique - Acquérir les bases du Machine Learning avec Python
■ <b>Modalités pédagogiques, techniques et d'encadrement :</b>	<ul style="list-style-type: none"><li>• Formation synchrone en présentiel et distanciel.</li><li>• Méthodologie basée sur l'Active Learning : 75 % de pratique minimum.</li><li>• Un PC par participant en présentiel, possibilité de mettre à disposition en bureau à distance un PC et l'environnement adéquat.</li><li>• Un formateur expert.</li></ul>
■ <b>Modalités d'évaluation :</b>	<ul style="list-style-type: none"><li>• Définition des besoins et attentes des apprenants en amont de la formation.</li><li>• Auto-positionnement à l'entrée et la sortie de la formation.</li><li>• Suivi continu par les formateurs durant les ateliers pratiques.</li><li>• Évaluation à chaud de l'adéquation au besoin professionnel des apprenants le dernier jour de formation.</li></ul>
■ <b>Sanction :</b>	Attestation de fin de formation mentionnant le résultat des acquis
■ <b>Référence :</b>	INT101603-F
■ <b>Note de satisfaction des participants:</b>	Pas de données disponibles
■ <b>Contacts :</b>	commercial@dawan.fr - 09 72 37 73 73

■ <b>Modalités d'accès :</b>	Possibilité de faire un devis en ligne ( <a href="http://www.dawan.fr">www.dawan.fr</a> , <a href="http://moncompteformation.gouv.fr">moncompteformation.gouv.fr</a> , <a href="http://maformation.fr">maformation.fr</a> , etc.) ou en appelant au standard.
■ <b>Délais d'accès :</b>	Variable selon le type de financement.
■ <b>Accessibilité :</b>	Si vous êtes en situation de handicap, nous sommes en mesure de vous accueillir, n'hésitez pas à nous contacter à <a href="mailto:referenthandicap@dawan.fr">referenthandicap@dawan.fr</a> , nous étudierons ensemble vos besoins

## Comprendre l'apprentissage automatique (Machine Learning)

Machine Learning : définition, contextes d'utilisation, phases (apprentissage, mise en production)

Enjeux et limites

Approches mathématiques et statistiques

Application et types de données : graphes, arbres, courbes, vecteurs de caractéristiques  
Qualification de la phase d'apprentissage : classification, régression, renforcement, non supervision

Panorama d'algorithmes : régressions (linéaire ou logistique), machines à vecteurs, réseaux de neurones, k plus proches voisins (KNN), boosting,...

Facteurs de pertinence et d'efficacité : nombre d'exemples, qualité des attributs, pourcentage des données renseignées ou manquantes, bruit

## Organiser un projet d'apprentissage

Définition du problème

Acquisition des données

Analyse et exploration des données

Préparation et nettoyage des données

Extraction de caractéristiques

Choix ou construction du modèle d'apprentissage

Entraînement, évaluation et optimisation

Test et vérification de surapprentissage Déploiement

## Découvrir des bibliothèques Python

NumPy : manipulation de matrices et fonctions

Pandas : lecture et manipulation de données

Jupyter et ses Notebook : utilisation de cahiers électroniques

## Implémenter des algorithmes d'apprentissage sur des données

Présentation d'une base de données

Régression linéaire (simple ou multiple) : étude, import, création et application d'un modèle, évaluation

Régression logistique : packages, création et application d'un modèle, évaluation et matrice de confusion

K plus proches voisins (KNN) : packages Python, Application et évaluation, amélioration

Support Vector Machine (SVM) : principe, découverte de la base de données, application et évaluation

Analyse en composante principale (PCA) : principe, création du modèle et visualisation des données en 2D

Decision Tree : principe, préparation des données, création du modèle et visualisation de l'arbre de décision

Random Forest : principe, implémentation en Python

## **Evaluer les modèles implémentés**

Ré-échantillonnage

Représentativité des données d'apprentissage  
Interprétation de la matrice de confusion  
Sensibilité et spécificité d'un test : Receiver Operating Characteristic (ROC) et Area under the ROC Curve (AUC)