
Programme de Formation

Data Science : maîtriser les fondamentaux pour débiter

Organisation

Durée : 14 heures

Mode d'organisation : Présentiel

Contenu pédagogique



Public visé

Développeurs, programmeurs, analystes de données ou statisticiens, ou toute personne disposant de bases en programmation et souhaitant acquérir les fondamentaux de la data science et de l'apprentissage automatique.



Objectifs pédagogiques

- Identifier les enjeux de la Data Science et de l'Intelligence Artificielle (IA).
- Schématiser le cycle complet d'un projet de Data Science.
- Appliquer les meilleures pratiques en matière de nettoyage et de préparation des données.
- Modéliser un problème de Data Science en utilisant des algorithmes d'apprentissage automatique.



Description

Jour 1

Matin

Introduction à la Data Science

Définition de la Data Science et distinction avec les statistiques et l'IA.

Applications et domaines d'intervention de la Data Science.

Outils et algorithmes de base pour la Data Science.

Cycle complet d'un projet Data Science : collecte des données, nettoyage, analyse, modélisation.

Après-midi

L'Apprentissage Automatique

Les différentes approches d'apprentissage automatique : supervisé, non supervisé, renforcé.

Rappel sur Python : structures de données, fonctions, POO (programmation orientée objet).

Librairies Python : NumPy, SciPy, Pandas pour le traitement des données et Matplotlib pour la visualisation.

Les algorithmes d'apprentissage automatique

Introduction à l'Analyse en Composantes Principales (ACP) et réduction de dimensionnalité.

Régression linéaire, multilinéaire et logistique avec régularisation.

Arbres de décision : classification supervisée et non supervisée.

KNN (K-Nearest Neighbors) et K-Means clustering.

Introduction aux SVM (Support Vector Machines) et aux réseaux de neurones.

Jour 2

Matin

Les essentiels de la Data Science

Cross-validation et utilisation des métriques d'évaluation.

Prise en charge de l'overfitting (surapprentissage) et de la gestion du biais vs variance.

L'importance du Feature Engineering pour créer des attributs pertinents.

La malédiction de la dimensionnalité et bonnes pratiques.

Après-midi

Data Science et Big Data

Traitement des données massives avec Spark MLlib.

Data visualisation avancée : exploration et présentation des résultats sous forme graphique.

Travaux pratiques

Création et amélioration d'un projet Data Science : Les apprenants travailleront sur un projet complet de Data Science en partant de l'exploration des données, du nettoyage et de la préparation des données, jusqu'à l'entraînement des modèles. Ils appliqueront l'ingénierie des attributs (Feature Engineering), amélioreront la qualité des données (gestion des valeurs manquantes, normalisation) et procéderont à un réapprentissage des modèles en optimisant leurs paramètres.



Prérequis

Connaissances de base en statistiques et pratique du langage Python.



Modalités pédagogiques

- Notre approche pédagogique repose sur une alternance entre apports théoriques opérationnels et mises en pratique directement issues du contexte professionnel des participants.
- Chaque thème traité s'appuie sur des études de cas concrets et des exercices pratiques, conçus spécifiquement à partir des situations réelles rencontrées dans l'environnement professionnel des stagiaires.
- Nos méthodes favorisent l'adaptation aux besoins de chaque participant afin d'optimiser leur engagement et leur apprentissage.
- L'animation privilégie une pédagogie active, favorisant les échanges, l'analyse réflexive des pratiques et une adaptation continue aux réalités professionnelles des participants.



Moyens et supports pédagogiques

- Des supports pédagogiques numériques adaptés au contexte professionnel sont remis aux participants.
- Dispositifs pédagogiques variés sont mobilisés tout au long de la formation (études de cas, exercices pratiques, travaux en sous-groupes, mises en situation, jeux pédagogiques) favorisant l'apprentissage actif, collaboratif et l'appropriation des contenus.



Modalités d'évaluation et de suivi

- Un QCM est réalisé en début et en fin de formation afin d'évaluer l'évolution des connaissances.
- Le suivi des stagiaires est assuré tout au long de la formation à travers des ateliers pratiques, des mises en situation et des échanges permettant d'évaluer leur participation et leur progression.

Une attestation individuelle de formation sera remise à chaque participant à l'issue de la formation.



Informations sur l'accessibilité

Si vous êtes en situation de handicap, ou si vous rencontrez une difficulté pouvant impacter votre participation à la formation, n'hésitez pas à nous contacter. Nous étudierons ensemble les aménagements ou adaptations pédagogiques, matérielles ou organisationnelles à mettre en place pour

vous accompagner dans les meilleures conditions.