

**NOUVEAUTE**

IA - CONDUIRE UN PROJET DE DATA SCIENCE ET DE MACHINE LEARNING

COMPÉTENCE PRINCIPALE VISÉE

- Acquérir la méthodologie et les outils d'un projet de data science

OBJECTIFS PÉDAGOGIQUES

- Découvrir et prendre en main Python pour des projets de Data Science
- Acquérir les concepts de base et la méthodologie d'un projet de Data Science
- Mettre en œuvre les outils nécessaires à la construction d'un projet de Data Science

PUBLIC

- Tout collaborateur en charge de mettre en œuvre un projet de Data Science

PRÉREQUIS

- Notion de programmation
- Niveau équivalent Bac+2 scientifique

CONTENU

PARTIE 1 - INTRODUCTION

- Introduction à la Data Science : présentation de l'histoire de la Data Science, des concepts de base, de cas d'étude et de la méthodologie
- Comprendre ce qu'est un dataset, les différents types de données manipulées en Data Science - transactionnel, graphe, ordonné - et acquérir des notions sur la qualité des données
- Rappels des traitements statistiques sur les données : moyenne, variance, écart-type, skewness, etc.

PARTIE 2 - FONDAMENTAUX DE PYTHON POUR LA DATA SCIENCE

- Concepts de base : types de données simples et composées, structure de contrôle, manipulation de fichiers
- Pourquoi Python pour la Data Science et découverte de l'écosystème de librairies, framework et outils
- Découverte de l'outil indispensable pour la manipulation de données en Python : Pandas
 - Concepts de base : Séries et DataFrame, lecture et écriture de données, sélection de données, qualité des données
 - Méthodes de visualisation de données
 - Calculs statistiques
 - Opération complexes : agrégation, pivot, jointure
 - Manipulation de séries temporelles

PARTIE 3 - MACHINE LEARNING

- Introduction aux activités du Machine Learning : clustering, prédiction, règles d'association, détection d'anomalie
- Apprentissage supervisé - régression : régression linéaire avec régularisation, notions de biais et variance, méthode de sélection des hyperparamètres, métriques

SESSIONS

L'ouverture de la session est conditionnée par un nombre minimum de participants.

DURÉE

7 jours (49 heures)

ÉQUIPE PÉDAGOGIQUE

L'équipe enseignante est composée de plusieurs enseignants-chercheurs de l'INSA Lyon, membres du laboratoire LIRIS, avec une activité de recherche reconnue en sciences des données. Cette équipe dispose d'une forte expertise technico-scientifique, obtenue via des nombreux projets de recherche partenariale, avec des grands groupes français mais également avec des start-up, PME ou ETI.


RENSEIGNEMENTS ET INSCRIPTION

Tel : +33 (0)4 72 43 83 93

Fax : +33 (0)4 72 44 34 24

mail : formation@insavalor.fr

Préinscription sur formation.insavalor.fr

 Accueil des personnes en situation de handicap nécessitant un besoin spécifique d'accompagnement : nous contacter à l'inscription



- Apprentissage supervisé - classification : régression logistique, arbre de décision, sous/sur-apprentissage, SVM, forêt aléatoire, ensemble learning, métriques
- Apprentissage non supervisé : clustering k-means et DBSCAN, clustering hiérarchique, méthodes d'analyse factorielles

MOYENS ET MÉTHODE PÉDAGOGIQUE

La démarche pédagogique proposée s'appuiera sur l'alternance d'apports théoriques, d'échanges techniques et d'illustrations autour de cas pédagogiques existants. En complément, elle intégrera de nombreuses mises en situation et de travaux pratiques

Entre les sessions de formation, des projets seront à réaliser par les stagiaires

ÉVALUATION ET RÉSULTATS

Évaluation des acquis de la formation

Évaluation des acquis des apprenants réalisée en fin de formation par un questionnaire ouvert contextualisé.

Taux de réussite

90.5% des apprenants ont acquis la compétence principale visée

Résultat obtenu pour 155 participants évalués ayant suivi une formation dans la thématique sur les 5 dernières années

Évaluation de la satisfaction

Évaluation du ressenti des participants en fin de formation (Niveau 1 KIRKPATRICK)

Résultats de l'évaluation

Le niveau de satisfaction globale est évalué à 4.3/5 par les participants.

Évaluations réalisées auprès des 192 participants ayant suivi une formation dans la thématique sur les 5 dernières années