

Datascience avec Python Scikitlearn

L'objectif de cette formation est de comprendre les utilisations de la datascience et l'intelligence artificielle, et de mettre en oeuvre concrètement le Machine Learning : applications de la Datascience, cycles d'un projet, algorithmes de ML, modélisation et mise en oeuvre avec ScikitLearn. *Pré-requis* : Python.



Modalité : à distance

Formule au choix : Live Training+ *ou bien*
Classe virtuelle Teams ou Zoom

Durée totale : 21 H (3 jours)

Projet Data Science : exemple : #5

Expression des besoins
Exploration des données avec Python
Visualisation avec Matplotlib
Nettoyage des données
Valeurs manquantes
Réduction de la dimension, ACP
Entraîner un modèles ML
Les métriques d'évaluation
Overfitting ou surapprentissage
Biais vs variance. Cross-validation
Réapprentissage
Bonnes pratiques

PLAN DETAILLE

Introduction à la Data Science : #1

Qu'est-ce que la Data Science ?
Statistiques, Data Science et IA
Champs d'application de la Data Science
Data Science : outils et algorithmes
Cycles d'un projet Data Science

L'apprentissage automatique : #2

Apprentissage supervisé
Apprentissage supervisé
Apprentissage Renforcé

Python pour la Datascience : #3

Syntaxe, structures de données
Fonctions, modules, librairie standard
Programmation orientée objet
Modules : NumPy, Pandas, Matplotlib, SciPy

Algorithmes du ML : #4

Analyse en composantes principales
Régression linéaire, multilinéaire
Régression polynomiale, logistique
Classification supervisée
Classification non supervisée
Régression avec arbres de décision
K-Nearest Neighbors
K-Means clustering
Réseaux de neurones artificiels

TP : mise en oeuvre ScikitLearn : #6

Extraction de caractéristiques et prétraitement
Réduction de dimension, ACP
Régression linéaire, multi-linéaire
Régression polynomiale, logistique
Classification supervisée
Classification non supervisée
Régression avec arbres de décision
K-means clustering
Réseaux de neurones artificiels