

Prérequis

- Python
- Bonnes connaissances de l'apprentissage automatique et de ses algorithmes

Programme (3j)

Formation sur 3 jours réparti à 50%/50% entre cours et pratique

Introduction (1h30)

- Les avantages de Python pour le calcul scientifique
- Présentation des librairies principales
- Les data sciences
- Installation des outils

Numpy (5h30)

- Introduction
- Les tableaux
- Création de tableaux
- Sélection de données
- Calculs et fonctions
- Filtres
- Association de tableaux
- Entrée / Sortie : formats de fichier standard

TP : Manipulation de numpy pour le calcul scientifique

Pandas (5h30)

- Introduction aux structures de données pandas (DataFrame et Series)
- Création de Série et de DataFrame
- Sélection des données
- Agrégation
- Lecture et sauvegarde de données dans des fichiers
- Fusion de DataFrame
- Manipulation de séries temporelles
- Utilisation d'une base de données

TP : Manipulation de Pandas pour données tabulaires et temporelles

Pytorch (5h30)

- Tenseurs
- Autograd
- Layers

- Optimisation
- Customisation

TP : Création d'un modèle Deep Learning avec Pytorch

PySpark (3h)

- Fonctionnement de Spark
- Interaction avec Spark en utilisant PySpark
- Manipulation de données
- Utilisation de MLlib

TP : Manipulation de données avec PySpark et utilisation de MLlib