

IA POUR LA RECHERCHE SCIENTIFIQUE

Jour 1 (6h) : Fondamentaux & Outils professionnels

🎯 **Objectif** : Comprendre les bases de l'IA et prendre en main les outils nécessaires.

Matin (3h)

- Introduction générale
 - Définition : IA, Machine Learning, Deep Learning
 - Domaines d'application scientifiques et industriels
 - Types d'apprentissage (supervisé, non supervisé, renforcement)
- Panorama méthodologique : **CRISP-DM** (vue d'ensemble)

Après-midi (3h)

- Installation et prise en main des outils :
 - Anaconda (environnements)
 - Jupyter Notebook (workflow interactif)
 - Scikit-learn (présentation rapide)
- Python scientifique :
 - **NumPy** : tableaux, opérations
 - **Pandas** : DataFrames, import/export CSV
- Atelier pratique : création d'un environnement, chargement d'un dataset simple, premières statistiques.

Jour 2 (6h) : Manipulation & Visualisation des données

🎯 **Objectifs:** Explorer, transformer et visualiser les données scientifiques.

Matin (3h)

- Pandas avancé :
 - Nettoyage, gestion des valeurs manquantes, doublons
 - Jointures, agrégations, filtrage avancé
- Data preprocessing : normalisation, feature engineering simple
- Atelier pratique : préparation d'un dataset "brut" (ex. country.csv, uforeports.csv).

Après-midi (3h)

- Visualisation avancée avec **Matplotlib**
 - Histogrammes, scatter plots, boxplots
 - Visualisation de corrélations (heatmap, matrices)
 - Mise en forme et personnalisation des graphiques
- Atelier pratique : réaliser un mini-rapport visuel sur un dataset scientifique.

Jour 3 (6h) : Algorithmes de Machine Learning

🎯 **Objectifs :** Découvrir, comprendre et appliquer les principaux algorithmes supervisés.

Matin (3h)

- Étapes d'un projet ML (CRISP-DM en détail) :
 - Préparation des données → Modélisation → Validation → Évaluation
- Algorithmes supervisés :
 - Régression linéaire et logistique
 - Classification (arbres de décision, k-NN, Random Forest)

- Évaluation des modèles : précision, rappel, F1-score, MAE, RMSE, R^2
- Atelier pratique : entraînement d'un modèle simple avec Scikit-learn.

Après-midi (3h)

- Introduction aux réseaux de neurones (perceptron, couches, activation)
- Utilisation de **Keras/TensorFlow** pour un petit réseau
- Comparaison de plusieurs modèles sur le même dataset
- Atelier pratique : classification simple avec un réseau de neurones.

Jour 4 (6h) : Mini-projet complet

🎯 **Objectif** : Mettre en œuvre un projet de bout en bout avec un dataset réel.

Matin (3h)

- Lancement du mini-projet (par groupe ou individuel)
 - Choix du dataset (santé, éducation, environnement, immobilier...)
 - Nettoyage et préparation des données
 - Visualisation et analyse exploratoire
- Travail guidé sur la modélisation (séparation train/test, sélection d'algorithmes).

Après-midi (3h)

- Construction et comparaison de plusieurs modèles :
 - Régression, classification, réseau de neurones
 - Visualisation des résultats et analyse des performances
- Présentation des résultats par les participants
- Conclusion & perspectives :
 - Limites des modèles et bonnes pratiques
 - Ressources pour progresser en IA et ML