

Programme complet de formation Data Science~~03 mois (12 semaines)

Orientation : Professionnalisation + Projets pratiques

Mois	Module principal	Objectif global
Mois 1	Outils & Fondamentaux de la Data Science	Maîtriser l'environnement, les outils, et la manipulation de données
Mois 2	Machine Learning	Appliquer les algorithmes d'apprentissage supervisé et non supervisé
Mois 3	Deep Learning & Projet de Fin de Formation	Créer des modèles avancés et livrer un projet complet de Data Science

DÉTAIL DU PROGRAMME

♦ **MODULE 1 : Environnement & Fondamentaux de la Data Science (Semaines 1–4)**

Semaine 1 : Mise en place & Git/GitHub

Objectif : savoir gérer ses projets et collaborer comme un vrai data scientist.

- Introduction à la ligne de commande
- Introduction à Git
 - Installation et configuration
 - Commandes de base : `git init`, `add`, `commit`, `status`, `log`
 - Branches et fusions
- Utilisation de GitHub
 - Création de repository distant
 - Connexion locale ↔ GitHub
 - Gestion de projets collaboratifs
 - Publication de notebooks Jupyter sur GitHub

- **Mini-projet :**

Créer un dépôt GitHub personnel + y publier un premier projet Python.

✖ Semaines 2–4 : Fondamentaux Data Science

Objectif : manipuler, analyser et visualiser efficacement les données.

1. Introduction à la Data Science

- Cycle de vie d'un projet data
-
- Cycle de vie d'un projet MLOps et IA
- Collecte, nettoyage, visualisation et interprétation
- Introduction aux statistiques descriptives

2. NumPy

- Tableaux multidimensionnels, slicing, broadcasting
- Fonctions mathématiques et agrégations

3. Pandas

- DataFrame, import/export (CSV, Excel, JSON)
- Nettoyage, filtrage, regroupement, fusion de données

4. Visualisation de données

- Matplotlib : graphiques de base
- Seaborn : visualisations avancées (heatmap, pairplot, boxplot, etc.)
- Plotly and Dash

5. Mini-projet pratique :

Analyse exploratoire d'un dataset réel

- Nettoyage, exploration, visualisation et interprétation

♦ **MODULE 2 : Machine Learning (Semaines 5–8)**

Objectif : Comprendre et appliquer les principaux algorithmes de machine learning avec Scikit-learn.

Contenu :

1. Concepts clés du Machine Learning
 - Supervised vs Unsupervised
 - Jeu de données d'entraînement/test, overfitting / underfitting, validation des modèles avec python....
2. **Apprentissage supervisé**
 - Régression linéaire et logistique
 - Arbres de décision, Random Forests, SVM, k-NN
3. **Apprentissage non supervisé**
 - K-Means, PCA, DBSCAN, HDBSCAN
 - Visualisation des clusters
4. **Évaluation et optimisation des modèles**
 - Cross-validation, GridSearchCV, métriques de performance
 - Normalisation, sélection de features
5. **Mini-projet pratique**

♦ **MODULE 3 : Deep Learning & NLP (Semaines 9–11)**

Objectif : Découvrir les réseaux de neurones et le traitement du langage naturel.

Contenu :

1. Introduction au Deep Learning
 - Concepts de neurones, couches, activation, perte
 - TensorFlow/Keras

2. CNN pour la vision par ordinateur
 - Reconnaissance d'images (MNIST, CIFAR-10)
3. Introduction au NLP
 - Prétraitement de texte (tokenization, TF-IDF)
 - Sentiment analysis
4. **Mini-projet pratique :**

Classification d'images ou analyse de sentiments, de tweets

♦ **MODULE 4 : Projet de Fin de Formation (Semaine 12)**

Objectif : Réaliser un projet complet de bout en bout, versionné sur GitHub.

Étapes :

- Choix du thème du projet
- Nettoyage, exploration, modélisation, visualisation
- Versioning complet du projet sur GitHub
- Présentation finale



Livrables :

- Notebook Jupyter documenté
- Repository GitHub propre et structuré
- Présentation PowerPoint du projet

Outils & Environnement

- **Langage** : Python
 - **Librairies** : NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, TensorFlow/Keras
 - **Outils** :
 - Jupyter Notebook / Google Colab
 - Git & GitHub
 - Power BI / Streamlit pour la visualisation finale
 - Kaggle, UCI Repository, data.gouv.bj pour les datasets
-

Évaluations

-  Mini-projet en fin de module
-  Évaluation continue (quiz + code review)