



Machine Learning Pratique avec Python et Scikit-learn

 DURÉE
4 jours (28h)

 RÉFÉRENCE
BSI536

 CATÉGORIE
**Machine Learning,
Data Mining et Deep
Learning**

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- ✓ Comprendre les fondamentaux du Machine Learning et leurs applications concrètes
- ✓ Manipuler et préparer efficacement des données pour l'entraînement de modèles
- ✓ Concevoir, entraîner et évaluer des modèles de régression, classification et clustering
- ✓ Optimiser les performances des modèles et comprendre les biais et limites
- ✓ Mettre en œuvre un workflow complet allant de l'exploration des données au déploiement simple

POUR QUI ?

- ✓ Développeurs, Data Scientists et Data Engineers souhaitant acquérir une pratique solide du Machine Learning
- ✓ Ingénieurs ou analystes de données ayant des bases en Python et désirant appliquer des techniques de ML à leurs projets



☰ Programme détaillé

1/ FONDAMENTAUX PYTHON POUR LA DATA SCIENCE

- Présentation des bibliothèques essentielles : NumPy, Pandas, Matplotlib, Seaborn
- Calcul matriciel et manipulation avancée de DataFrames
- Nettoyage et préparation des données : gestion des valeurs manquantes, types de données, encodage
- Visualisation et exploration : histogrammes, heatmaps, boxplots, pairplots
- Mise en pratique : analyse exploratoire sur un dataset réel (Titanic, Bike Sharing)

2/ INTRODUCTION AU MACHINE LEARNING SUPERVISÉ

- Concepts clés : features, labels, apprentissage supervisé vs non supervisé
- Les principaux algorithmes de régression et classification
- Utilisation de scikit-learn : entraînement, prédiction, métriques (accuracy, précision, recall, RMSE)
- Automatisation avec les pipelines : preprocessing + modèle
- Mise en pratique : régression linéaire et logistique, k-NN, arbres de décision

3/ TECHNIQUES AVANCÉES ET MACHINE LEARNING NON SUPERVISÉ

- Les modèles d'ensemble : Random Forest, Gradient Boosting
- Réduction de dimensionnalité : PCA, t-SNE pour visualisation
- Clustering : k-means, DBSCAN, hiérarchique
- Optimisation des hyperparamètres : GridSearchCV, RandomizedSearchCV
- Mise en pratique : choix du bon modèle pour un problème donné + comparaison des résultats

4/ OPTIMISATION, INTERPRÉTATION ET DÉPLOIEMENT

- Comprendre le surapprentissage et régularisation (L1, L2)
- Explicabilité des modèles : SHAP, Feature Importance

- Bonnes pratiques d'évaluation : train/test split, validation croisée
- Introduction au déploiement : sérialisation avec joblib/pickle, mise en place d'un service simple avec Flask
- Mise en pratique : projet de bout en bout (de la préparation des données au déploiement d'un modèle simple)

Approche pédagogique

- ✓ Support Ecrit et Projection
- ✓ Exposés Interactifs, Podcasts et Vidéos
- ✓ Brainstorming et Jeux de Rôle
- ✓ Cas Pratiques et Labs inclus pour leur impact opérationnel
- ✓ Test de Validation des Acquis des Connaissances

Prochaines dates programmées

📅 30 Juin au 03 Juil. 2026

📍 Casablanca - Maroc

📅 25 au 28 Août 2026

📍 Casablanca - Maroc

📅 20 au 23 Oct. 2026

📍 Casablanca - Maroc

📅 Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Réservation & Renseignements

📞 **Téléphone** : +212 522 247 210

✉ **Email** : contact@innov-systems.com

🌐 **Web** : <https://www.innov-systems.com>

▼
Scannez pour accéder
à la fiche en ligne