

## MACHINE LEARNING SOUS PYTHON ET R

**Durée**

**4 jours**

**Référence Formation**

**4-PY-MLPR**

### Objectifs

Utiliser R et Python dans le cadre de modèles d'apprentissage et de fonctions algorithmiques

### Participants

Cette formation s'adresse au profil métier orienté Data Scientist

### Pré-requis

Bonnes bases de statistiques et de data mining

### Moyens pédagogiques

Accueil des stagiaires dans une salle dédiée à la formation équipée d'un vidéo projecteur, tableau blanc et paperboard ainsi qu'un ordinateur par participant pour les formations informatiques.

Positionnement préalable oral ou écrit sous forme de tests d'évaluation, feuille de présence signée en demi-journée, évaluation des acquis tout au long de la formation.

En fin de stage : QCM, exercices pratiques ou mises en situation professionnelle, questionnaire de satisfaction, attestation de stage, support de cours remis à chaque participant.

Formateur expert dans son domaine d'intervention

Apports théoriques et exercices pratiques du formateur

Utilisation de cas concrets issus de l'expérience professionnelle des participants

Réflexion de groupe et travail d'échanges avec les participants

Pour les formations à distance : Classe virtuelle organisée principalement avec l'outil ZOOM. Assistance technique et pédagogique : envoi des coordonnées du formateur par mail avant le début de la formation pour accompagner le bénéficiaire dans le déroulement de son parcours à distance.

## PROGRAMME

### Généralités

- Présentation et installation du logiciel R
- GUIs et IDEs
- Prise en main et premiers pas
- Fonctions et packages
- Utilisation des aides

### Rappel des concepts de base du Data Mining

- Introduction à R - Concepts de base
- Syntaxe
- Type de données
- Import/Export de données

- Manipulation des librairies

### **Manipulation des données**

- Rappels des grandeurs statistiques de bases
- Calcul des grandeurs statistiques de bases avec R
- Aperçu des types de visualisation courants en Data Mining
- Fonctions et librairies de base de visualisation avec R

### **Data Preprocessing**

- Data Cleaning
- Data Reduction (analyse en composantes principales)
- Data Preprocessing et ACP avec R

### **Règles d'association**

- Rappels théoriques de l'algorithme Apriori
- Application et exploitation des règles d'association avec R

### **Classification et régression**

- Règles bayésiennes naïves
- Rappels théoriques
- Initiation à la librairie "e1071" de R
- Arbre de décision CART
- Rappels théoriques
- Initiation à la librairie "party" de R
- Arbre de décision RandomForest
- Rappels théoriques
- Initiation à la librairie "randomForest" de R
- Régression linéaire
- Rappels théoriques
- Initiation à la fonction lm() de R

### **Clustering**

- Centroid-based clustering
- Algorithme du K-means
- Fonction kmeans de R
- Hierarchical clustering
- Concepts théoriques
- Fonction hclust de R
- Density-based clustering
- Concepts théoriques
- Application dans R

### **Outils de support à R**

- Data preprocessing avec Dataiku
- Utilisation de R dans Dataiku



**CAP ÉLAN FORMATION**

[www.capelanformation.fr](http://www.capelanformation.fr) - Tél : 04.86.01.20.50

Mail : [contact@capelanformation.fr](mailto:contact@capelanformation.fr)

Organisme enregistré sous le N° 76 34 0908834

version 2025