

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

Réflexion de groupe et apports théoriques du formateur

Travail d'échange avec les participants sous forme de réunion-discussion

Utilisation de cas concrets issus de l'expérience professionnelle

Validation des acquis par des questionnaires, des tests d'évaluation, des mises en situation et des jeux pédagogiques.

Remise d'un support de cours.

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

CAP ÉLAN FORMATION

www.capelanformation.fr - Tél : 04.86.01.20.50

Mail : contact@capelanformation.fr

Organisme enregistré sous le N° 76 34 0908834

version 2024

- 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 - Tél : 04.86.01.20.50

Mail : contact@capelanformation.fr

Organisme enregistré sous le N° 76 34 0908834
version 2024