Big Data Analytics avec R

Lien : https://innov-systems.com/formation/big-data-analytics-avec-r

DURÉE
5 jours (35h)

RÉFÉRENCE
BSI97

CATÉGORIE
Big Data - Techniques D'analyse

OBJECTIFS DE LA FORMATION

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

- Savoir installer R
- Comprendre le principe de la modélisation statistique
- Comprendre comment manipuler des données avec R
- Savoir choisir et réaliser de façon autonome des analyses factorielles et classifications

POUR QUI ?

- Ingénieurs
- Analystes
- Data Analysts
- Toute personne intéressée par l'analyse statistique avec R

Innov Systems

Programme détaillé

1 / Introduction générale

- Présentation du logiciel R
- Installation du logiciel R et packages

2 / Types de données dans R

- Manipulations de scalaires (réels)
- Manipulations de nombres complexes
- Manipulations de variables
- Manipulations de vecteurs
- Manipulations de matrices
- Manipulations de textes
- Manipulations de dates et de durées

3 / Les bases d'utilisation du logiciel R

- Importation et exportation de données
- Manipulation de données
- Tracer des courbes et des graphiques

4 / Analyse en composantes

- Analyse en Composantes Principales
- Analyse du nuage des individus
- Propriétés des composantes principales
- Analyse du nuage des variables
- Interprétation des résultats
- Analyse des correspondances

Innov Systems

- Nuages des profils-lignes et des profils-colonnes, la distance du khi-deux
- Analyse des nuages des profils
- Liens entre les analyses
- Interprétation des résultats
- Analyse des Correspondances Multiples
- Analyse des correspondances sur le tableau disjonctif complet
- Propriétés de l'ACM
- Analyse du tableau de Burt
- Analyse Factorielle pour Données Mixtes
- Classification
- Classification ascendante hiérarchique : principe, stratégies d'agrégation, aides à l'interprétation
- Méthodes de partitionnement : les méthodes des centres mobiles, les nuées dynamiques
- Classification mixte

5 / La modélisation

- Construction d'un modèle (étapes)
- Algorithmes supervisés et non supervisés
- La régression ou la classification

6 / Procédures d'évaluation de modèles

- Les techniques de ré-échantillonnage
- Test de représentativité des données d'apprentissage
- Mesures de performance des modèles prédictifs
- Matrice de confusion, de coût et la courbe ROC et AUC

7 / Les algorithmes non supervisés

- Le clustering hiérarchique
- Le clustering non hiérarchique
- Les approches mixtes

Inov Systems

8 / Les algorithmes supervisés

- Le principe de régression linéaire univariée
- La régression multivariée
- La régression polynomiale
- La régression régularisée
- Le Naive Bayes
- La régression logistique

9 / Analyse de données textuelles

- Collecte et prétraitement des données textuelles
- Extraction d'entités primaires, d'entités nommées et résolution référentielle
- Étiquetage grammatical, analyse syntaxique, analyse sémantique
- Lemmatisation
- Représentation vectorielle des textes
- Pondération TF-IDF
- Word2Vec

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

Innov Systems

Prochaines dates programmées

15 au 19 Fév. 2027	Présentiel - Casablanca
12 au 16 Avr. 2027	Présentiel - Casablanca
07 au 11 Juin 2027	Présentiel - Casablanca
02 au 06 Août 2027	Présentiel - Casablanca
27 Sep. au 01 Oct. 2027	Présentiel - Casablanca

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

Innov Systems — Tous droits réservés

Innov Systems