

## Machine Learning sous Python et R



### OBJECTIFS PÉDAGOGIQUES

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



### PUBLIC CONCERNÉ

- Cette formation s'adresse au profil métier orienté Data Scientist



### PRÉREQUIS

- 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
- Alternance entre apports théoriques et exercices pratiques (en moyenne 30 à 50%)
- Remise d'un support de cours.



### MODALITÉS D'ÉVALUATION

- Feuille de présence signée en demi-journée,
- Evaluation des acquis tout au long de la formation,
- Questionnaire de satisfaction,
- Attestation de stage à chaque apprenant,
- Positionnement préalable oral ou écrit,
- Evaluation formative tout au long de la formation,
- Evaluation sommative faite par le formateur ou à l'aide des certifications disponibles



### MOYENS TECHNIQUES EN PRÉSENTIEL

- Accueil des stagiaires dans une salle dédiée à la formation équipée à minima d'un vidéo projecteur et d'un tableau blanc et/ou paperboard.
- Pour les formations nécessitant un ordinateur, un PC est mis à disposition de chaque participant.



### MOYENS TECHNIQUES EN DISTANCIEL

- A l'aide d'un logiciel (Teams, Zoom...), d'un micro et éventuellement d'une caméra les apprenants interagissent et communiquent entre eux et avec le formateur.
- Sessions organisées en inter comme en intra entreprise.
- L'accès à l'environnement d'apprentissage ainsi qu'aux preuves de suivi et d'assiduité (émargement, évaluation) est assuré.
- Pour toute question avant et pendant le parcours, assistance technique à disposition au 04 67 13 45 45.



### ORGANISATION

- Délai d'accès : 5 jours ouvrés (délai variable en fonction du financeur)
- Les cours ont lieu de 9h à 12h30 et de 13h30 à 17h



### ACCESSIBILITÉ

- Les personnes en situation de handicap sont invitées à nous contacter directement, afin d'étudier ensemble les possibilités de suivre la formation.
- Pour tout renseignement, notre référent handicap reste à votre disposition : mteyssedou@ait.fr



### PROFIL FORMATEUR

- Nos formateurs sont des experts dans leurs domaines d'intervention
- Leur expérience de terrain et leurs qualités pédagogiques constituent un gage de qualité.



### CERTIFICATION POSSIBLE

- Aucune

# Machine Learning sous Python et R

## 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