

INITIATION AU TEXT MINING

Durée & Format

Durée : 14 heures

Format disponible : Présentiel ou Distanciel

Tarifs

Tarif public intra : Consulter notre [site Internet](#)

Tarif public inter : Consulter notre [site Internet](#)

Public visé

- Toute personne souhaitant découvrir le Text mining et sa mise en application sous R

Prérequis

- L'apprenant devra avoir des notions de bases en statistique (statistique descriptive, estimation), en Data Mining (apprentissage supervisé, non supervisé, ACP) ainsi que des notions de R.

Objectifs pédagogiques

A l'issue de cette formation, l'apprenant sera capable de :

- Importer une base de données textuelles sous R
- Représenter les données textuelles sous un format approprié
- Préparer les données en effectuant les principaux pré-traitements spécifiques aux données textuelles
- Visualiser les données textuelles
- Effectuer une recherche d'informations
- Effectuer une classification de documents (apprentissage supervisé)
- Extraire des thématiques d'une base de données textuelles (apprentissage non supervisé)

Programme et déroulement

Jour 1 – Matin

Introduction au Text Mining

- Contexte
- Exemple de données
- Application du Text Mining

Particularité des données textuelles

- Multiplicité des représentations
- Fléau de la dimension
- Subjectivité des données

Jour 1 – Après-midi

Représentations des données textuelles

- Sac de mots
- Vectorisation

Pré-traitements

ARKESYS.NET

70 rue Bergson – 42000 Saint-Étienne
04 28 95 15 82 - info@arkesys.fr
SAS AU CAPITAL DE 60 000 € - SIRET : 501 033 609 00030
ETS secondaire : 58 avenue Debourg - 69007 Lyon - SIRET : 501 033 609 00048
ETS secondaire : 293 route de la Seyne - 83190 Ollioules - SIRET : 501 033 609 00055



- Tokenization
- Nettoyage (majuscule, nombre, ponctuations...)
- Stopwords
- Stemming

Jour 2 – Matin

Visualisation

- Histogramme
- Nuage de mots

Recherche d'informations

- Comparaison de textes
- Pondération des termes (TF-IDF...)
- Text ranking

Jour 2 – Après-midi

Application du text mining

- Recherche d'information
- Classification de documents
- Analyse de sentiments
- Classification de documents
- Extraction de thématiques avec le modèle LDA

Certification

Consulter le programme sur notre [site internet](#) pour identifier si cette formation est certifiante.

Modalités pédagogiques

- Explications théoriques suivies de pratiques guidées puis de mises en autonomie.
- Exercices autonomes et réguliers pour assurer l'assimilation

Moyens et supports pédagogiques

Votre formation a lieu en présentiel :

- 1 vidéoprojecteur par salle
- 1 ordinateur

Votre formation se déroule à distance avec :

- 1 ordinateur
- 1 connexion Internet
- 1 adresse e-mail valide
- 1 équipement audio (micro et enceintes ou casque)
- 1 Webcam (facultatif – dans l'idéal)
- 1 deuxième écran (facultatif – dans l'idéal)

Votre formation se déroule sur notre plate-forme de formation avec :

- 1 ordinateur
- 1 connexion Internet
- 1 adresse e-mail valide
- 1 équipement audio (micro et enceintes ou casque - facultatif)

Support stagiaire :

À l'issue de la formation, les exercices et travaux pratiques réalisés, leurs corrigés ainsi qu'un support de cours dématérialisé sera fourni à chaque stagiaire par e-mail ou via la plate-forme FOAD.

Evaluation et suivi

Les objectifs pédagogiques sont évalués et suivis grâce à différentes **méthodes adaptées aux acquisitions de compétences visées**. Nos formateurs réalisent ces évaluations **tout au long de la formation**, que ce soit pendant les séances synchrones ou asynchrones. Voici une liste non exhaustive des méthodes d'évaluation pouvant être utilisées en formation :

- **Questionnaire** de vérification de connaissances (Quiz)
- Réalisation de **Travaux Pratiques** : production ou amélioration d'un fichier
- **Mises en situation** et grilles d'analyse
- **Serious Game** : jeu de rôles et analyse
- **Activités interactives** à travers l'utilisation d'une plate-forme connectée
- ...

Toutes nos formations intègrent **une auto-évaluation** via notre Extranet Stagiaire au début et à la fin de chaque action de formation. Cet outil offre à chacun la possibilité de mesurer sa progression par rapport aux objectifs pédagogiques visés et leurs atteintes.

Profil formateur

Nos formateurs sont certifiés à l'issue d'un parcours organisé par nos soins. Ils bénéficient d'un suivi de maintien et d'évolution de leurs compétences aussi bien au niveau technique que pédagogique. Chacun de nos formateurs a bénéficié d'une formation spécifique à l'animation de classe virtuelle et à l'utilisation des solutions de formation à distance du Groupe ARKESYS.

Support stagiaire

À l'issue de la formation, les exercices et travaux pratiques réalisés, leurs corrigés ainsi qu'un support de cours dématérialisé sera fourni à chaque stagiaire via son extranet stagiaire.

Accessible à tous

ARKESYS.NET

70 rue Bergson – 42000 Saint-Étienne

04 28 95 15 82 - info@arkesys.fr

SAS AU CAPITAL DE 60 000 € - SIRET : 501 033 609 00030

ETS secondaire : 58 avenue Debourg - 69007 Lyon - SIRET : 501 033 609 00048

ETS secondaire : 293 route de la Seyne - 83190 Ollioules - SIRET : 501 033 609 00055

Cette formation est accessible à toute personne en situation de handicap. Notre référent handicap prendra contact avec les stagiaires concernés pour adapter l'animation à leurs besoins et rendre l'apprentissage accessible à tous. Enfin, nos centres de formation sont accessibles aux personnes à mobilité réduite. [En savoir plus](#)