

Technologies de Text Mining

Le traitement automatique du langage ou text mining permet d'analyser du texte en combinant les outils divers comme le machine learning (IA), la linguistique et la sémantique, les statistiques et l'informatique. Dans cette formation, vous allez apprendre ce qu'est le text mining, vous repérer dans le périmètre technologique et mettre en œuvre ces technologies via des outils adaptés et des cas pratiques rencontrés en entreprise. **Pré-requis** : Python, IA et analyse de données.



Modalité : à distance

Formule au choix : Live Training+ *ou bien*
Classe virtuelle Teams ou Zoom

Durée totale : 14 H (2 jours)

PLAN DETAILLE

Introduction au Text Mining : #1

Qu'est-ce que le Text Mining?
Comment fonction le Text Mining?
Approches statistique et sémantique
Quels problèmes résoudre avec le TM?

IA et fouille de données : #2

Rappels IA et machine learning
Les étapes de la fouille de données
Les types de tâches
Exemples d'application

Outils et traitements en Text Mining : #3

Les langages utilisés
Outils Python pour le Text Mining
Les différents traitements
Le stemming. Lemmatization
Stemming vs lemmatization
Autres traitements

NLP Deep Learning : #6

Introduction à NLP
Librairies NLP. Librairies Deep Learning
Pratique du NLP. Recurrent Neural Networks RNN
Modèles RNN. Travaux pratiques

Classification non supervisée : #4

Similarité des textes
Distances : Hamming, Manhattan, etc.
Analyse de la similarité des documents.
Algorithmes de classification.
Travaux pratiques

Classification supervisée : #5

Prétraitement, normalisation des données.
Modèles de classification. Régression logistique. Support Vector Machines. Random Forest.
Gradient Boosting Machines.
Évaluation des modèles. Travaux pratiques