

Formation Text Mining

Réf. SF33911

Maîtrisez l'art de l'analyse textuelle pour transformer les données en connaissances pratiques avec notre formation de 3 jours !

3 j Durée	2 700 € HT Inter / participant	5 550 € HT Intra (≤12 pers.)	0 Sessions à venir
--------------	-----------------------------------	---------------------------------	-----------------------

Objectifs pédagogiques

- ✓ Identifier les fondements théoriques et les applications pratiques du Text Mining
- ✓ Appliquer les méthodes traditionnelles et les approches modernes d'analyse
- ✓ Mettre en œuvre des techniques de prétraitement et de feature engineering
- ✓ Évaluer la similarité entre textes et effectuer la classification non supervisée
- ✓ Implémenter des modèles de classification supervisée pour catégoriser les documents
- ✓ Appliquer des techniques avancées de NLP et de Deep Learning pour des tâches spécifiques en traitement de texte

Public visé & prérequis

Public visé Chefs de projet, Ingénieurs Big Data, Tous

Prérequis Connaissances en statistiques, traitement de données, Machine Learning et Deep Learning.

Programme détaillé

Introduction à la formation Text Mining

Objectifs et contexte de la formation
Importance du Text Mining dans le contexte actuel
Aperçu du programme

Fondements du Text Mining

Définition du Text Mining
Principes de base et évolution
Applications pratiques et enjeux dans divers secteurs

Processus et Approches Traditionnelles

Récupération des données textuelles avec les API
Récupération et exploration du corpus de textes
Approches traditionnelles telles que Bag of Words (BoW) et N-Grams
Modèles de langue probabilistes et méthodes statistiques
Prétraitement des données textuelles (suppression des caractères accentués, stemming, lemmatization)

Activités pratiques :

Comparaison des résultats obtenus par les différentes approches sur des corpus variés

Feature Engineering pour la représentation de texte

Syntaxe et structure de textes
Vectorisation de mots et de documents
Techniques de word embeddings (Word2Vec, GloVe)
Approches basées sur les modèles de plongement de phrases
Utilisation du modèle TF-IDF, Transformer et Vectorizer

Construction de matrices de termes-document
Sélection des caractéristiques pertinentes pour l'analyse textuelle

Activités pratiques :

Création et utilisation de vecteurs de mots pour améliorer la représentation des textes

Similarité des textes et classification non supervisée

Mesures de similarité textuelle
Regroupement de documents (Clustering)
Techniques de réduction de dimension pour l'analyse de similarité
Identification des thèmes latents dans les corpus
Application de méthodes non supervisées à des jeux de données réels
Algorithmes de classification non supervisée

Activités pratiques :

Clustering de documents et identification de thèmes pertinents dans un ensemble de données spécifique

Classification supervisée du texte

Méthodes classiques de classification supervisée
Prétraitement et normalisation des données
Sélection des caractéristiques pour la classification
Explication des algorithmes tels que Multinomial Naïve Bayes, Régression logistique, Support Vector Machines, Random Forest, Gradient Boosting Machines
Évaluation des modèles de classification textuelle

Activités pratiques :

Construction et évaluation d'un modèle de classification pour la catégorisation de documents

Natural Language Processing (NLP) et Deep Learning

Introduction au traitement automatique du langage naturel (NLP)
Réseaux de neurones pour le traitement du langage naturel : NLTK, TextBlob, SpaCy, Gensim, Pattern, Stanford CoreNLP
Architectures de réseaux récurrents et réseaux neuronaux convolutifs
Applications avancées du Deep Learning en NLP : Theano, TensorFlow, Keras
RNN, Long Short-Term Memory, modèles bidirectionnels et Sequence-to-Sequence
Modèles RNN pour questions et réponses
Perspectives futures et tendances dans le Text Mining

Activités pratiques :

Application de modèles de Deep Learning pour des tâches spécifiques en traitement de texte

Pourquoi choisir Sparks Formation ?

20 ans d'expertise IT, des formateurs experts et des sessions garanties contre l'annulation.

20 ans
d'expertise

1 068+
formations

11 178
apprenants formés/an

+67
Net Promoter Score

Avis des participants

4,9/5 — 19 avis

Anthony G. — Formation très bien organisée, très claire Alison explique très clairement et est très à l'écoute, avec toujours une approche pratique essentielle. Je n'avais jamais fait de Python et...

François H. — François-Marie est un excellent formateur ! Un grand merci à lui d'avoir été aussi pédagogue et clair dans ses explications du début à la fin de la formation,...

Ils nous font confiance

Cette année notre filière Data Science & Machine Learning a accompagné plus de 55 entreprises dont :

SOPRA STERIA RTE AMADEUS SAS CEGEDIM LBC FRANCE SNCF - E.SNCF SOLUTIONS CSTB GERS SA ADENTIS ADVENTIEL +45 autres

Financement & certifications

Cette formation est éligible aux financements OPCO et FNE-Formation. Nos équipes vous accompagnent dans le montage de votre dossier.



Réservez votre place dès maintenant

20 ans d'expertise · 1 000+ formations · NPS +67 · réponse sous 24h ouvrées

0 805 950 800

demande@sparks-formation.com