Text Mining : techniques et applications

Référence et durée

TM_FD
1 jours, 7 heures

Objectif

- Acquérir la méthodologie d'étude à mettre en œuvre pour étudier les données textuelles

Public

Data Scientists.

Prérequis

Avoir quelques connaissances sur le Machine Learning.

Programme

- Introduction et domaines d'application
- Représentation des données
  - Loi de Zipf
  - Attributs
  - Valeurs et pondération : sacs de mots, vecteurs
  - Mesures de distance et similarité
- Recherche d'information (RI)Re
  - Indexation
  - RI booléenne
  - RI vectorielle
  - Page Rank
- Classification
  - Classification supervisée
  - Classification non supervisée
- Annotation
  - HMM
  - CRF
- Exemples d'applications
  - Moteur de recherche
  - Analyse de sentiments
- Démonstrations