LINCOLN Light up your data

Machine learning : algorithmes et applications

Référence et durée

ML_FD
1 jour, 7 heures

Objectifs

Avoir une vision transverse des principaux algorithmes et méthodes d'analyse prédictive supervisée.

Public et prérequis

Cette formation s'adresse à un public d'analystes, data miners, chargés d'études statistiques, directeurs d'études. Des connaissances statistiques et informatiques sont préférables.

Moyens pédagogiques

Alternance de présentations théoriques et de démonstrations.

Programme détaillé

- Introduction
  - Algorithmes, vocabulaire
  - Démarche d'apprentissage : feature extraction (décomposition de matrices, n-grams, hashing, tf-idf
  - Démarche d'apprentissage : échantillonnage, validation, cross-validation, grid search, feature selection, metriques de validation (mape, AUC, …), biais-variance, erreur d'apprentissage, erreur réelle
  - Démarche d'apprentissage : les algorithmes
    - Les méthodes supervisées : knn, SVM, Réseaux de neurones, réseau bayésien naif, régression pénalisée, boosting, random forest
    - Les méthodes non supervisées : Clustering, Règles d'associations et de séquences, Traitement du texte (text mining)
- Démonstrations d'algorithmes non supervisés
  - Analyse de données textuelles : feature extraction, analyse latente
- Démonstrations d'algorithmes supervisés
  - Mise en œuvre d'un modèle (apprentissage supervisé d'une cible binaire ou continue) à partir d'un grand nombre de variables
  - Métriques d'Evaluation et comparaison de modèles