

R5.02 Data mining

Composante
Institut universitaire de technologie de Poitiers-Châtelleraut-Niort

Présentation

Description

L'objectif de cette ressource est de comprendre les principales méthodes d'apprentissage automatique supervisé, connaître leurs implémentations et savoir en interpréter les résultats.

Contenus :

- Définir les objectifs des méthodes apprentissage supervisé
- Déroulement d'une procédure d'apprentissage automatique supervisé (échantillons d'apprentissage, développement, test, ...)
- Arbres de décisions CART (classification et régression)
- Analyse discriminante linéaire dont le cas du Classifieur Bayésien naïf
- Méthode des K plus proches voisins
- Applications sur variable-cible binaire ou multi-classes.
- Traitement des variables qualitatives. Cas de fichiers mixtes.
- Évaluation et comparaison des modèles (mesures d'évaluation, validation croisée, bootstrap)

L'utilisation d'un logiciel d'analyse statistique pour la mise en œuvre des méthodes est indispensable. Cette ressource s'appuiera sur la ressource « Classification automatique ».

Heures d'enseignement

CM	CM	9h
TD	TD	21h