

Master Mathématiques appliquées, statistique

Classification

Informations

Composante : Faculté des Sciences

Langue(s) d'enseignement

Français

Contenu

Classification non supervisée : Distances et dissimilarités usuelles : euclidienne, Mahalanobis, L1, khi-deux. Classification par partition : algorithme des centres mobiles. Classification ascendante hiérarchique, stratégies d'agrégation : lien moyen et Ward. Statistiques permettant le choix d'une partition, interprétation des partitions (valeurs tests)

(Optionnel : introduction à DBSCAN et/ou au modèle de mélange gaussien)

Classification supervisée : Régression logistique binaire : notion de cote et rapport de cote. Estimateurs du maximum de vraisemblance : tests d'hypothèses et intervalles de confiance. Critères d'évaluation de l'ajustement. Approche décisionnelle : règle d'affectation, sensibilité, spécificité et courbe ROC. Estimation du taux d'erreur de prédiction par validation croisée.

Compétences à acquérir

- Mobiliser un algorithme de classification pour partitionner un ensemble d'observations
- Évaluer les qualités de la partition obtenue
- Mobiliser un modèle de régression généralisée pour expliquer une variable binaire.
- Interpréter les estimations du modèle et d'utiliser le modèle dans le cadre d'une approche de type décisionnel

Bibliographie, lectures recommandées

Gilbert Saporta (2006) Probabilités, Analyse des données et Statistique, Technip.

J.P. Nakache, J. Confais (2003) Statistique Explicative Appliquée, Technip (Partie 2)

Lebart, Piron, Morineau (2006). Statistique exploratoire multidimensionnelle. 4ème édition. Dunod

Ricco Rakotomalala. Pratique de la Régression Logistique - Régression Logistique Binaire et Polytomique

VOLUME HORAIRE

- Volume total: 24 heures
- Cours magistraux: 12 heures
- Travaux pratiques: 12 heures

Codes Apogée

- SMSBU32C [ELP]

Pour plus d'informations

[Aller sur le site de l'offre de formation...](#)



Dernière modification le 07/10/2024