09
MAR.
2016

de 12h15 à 13h30

SÉMINAIRE

Clustering de données ordinales

Julien JACQUES
Laboratoire ERIC - Université de lyon

Du fait du manque d'outils spécifiques, les données ordinales sont souvent assimilées soit à des données nominales, oubliant la notion d'ordre, soit à des données quantitatives, introduisant artificiellement une notion de distance entre modalités. Dans le but d'éviter l'utilisation d'une de ces deux solutions extrêmes, nous proposons une nouvelle distribution de probabilité pour données ordinales, paramétrée par un paramètre de position et un paramètre de précision. Cette distribution est ensuite utilisée pour définir un algorithme de clustering spécifique aux données ordinales, permettant de prendre en compte les données multivariées et potentiellement manquantes. Cet algorithme a été implémenté dans des solutions logiciels (package R, logiciel en ligne SaaS), qui seront utilisés pour une démonstration sur données réelles.Contact : Mme Mariethé CHAUMEIL Inscription gratuite mais obligatoire au plus tard le lundi 29 février 2016 Courriel : mariethe.chaumeil@chu-lyon.fr