

Note de lecture



Approche pragmatique de la classification.
par Josiane Confais et Jean-Pierre Nakache
Technip, 2004, 272 p., 49 €.

Les techniques de classification figurent parmi les méthodologies les plus sollicitées par des approches exploratoires tant pour les sciences du vivant que pour celles de la société. Les technologies pour l'extraction d'information s'appuient désormais largement sur de nouvelles approches algorithmiques développées pour la fouille d'entrepôts de données recelant des ensembles d'objets dont la cardinalité est très grande.

L'ouvrage que nous proposent Jean-Pierre Nakache et Josiane Confais, est le fruit de leurs expériences de formateurs, tant en formation initiale au sein de l'Institut de Statistique de l'Université Pierre et Marie Curie (ISUP), qu'en formation continuée dans le cadre des stages organisés par le CISIA-CERESTA. L'approche de l'ouvrage est effectivement pragmatique, s'appuyant sur de nombreux exemples, les logiciels les plus couramment utilisés en la matière ainsi qu'une bibliographie très riche à la fois au plan méthodologique mais également au plan applicatif. Les auteurs passent bien sûr en revue les techniques classiques proposées par les logiciels commerciaux (SAS, Splus, SPSS, Statistica) ou libres (R), mais donnent également un aperçu sur les développements les plus récents, par exemple certains algorithmes de classification conduisant à des classes non convexes de forme arbitraire (*shrinking*).

Après un chapitre introductif consacré aux mesures de la ressemblance ou de la dissemblance entre objets taxonomiques, la structure de l'ouvrage reprend la distinction communément admise pour les techniques de classification entre méthodes hiérarchiques et méthodes de partitionnement. La partie de l'ouvrage consacrée aux méthodes hiérarchiques traite de façon approfondie les algorithmes de type agglomératif (classification ascendante