

Classification Conceptuelle avec Généralisation par Intervalles

Paula Brito*, Géraldine Polaillon**

*Faculdade de Economia & LIAAD-INESC Porto LA, Universidade do Porto
Rua Dr. Roberto Frias, 4200-464 Porto, Portugal
mpbrito@fep.up.pt , <http://www.fep.up.pt/docentes/mpbrito>

**SUPELEC Science des Systèmes (E3S) - Département Informatique
Plateau de Moulon, 3 rue Joliot Curie, 91192 Gif-sur-Yvette cedex, France
geraldine.polaillon@supelec.fr

Résumé

Nous nous intéressons aux méthodes de classification hiérarchique ou pyramidale, où chaque classe formée correspond à un concept, i.e. une paire (extension, intension), considérant des données décrites par des variables quantitatives à valeurs réelles ou intervalles, ordinales et/ou prenant la forme de distribution de probabilités/fréquences sur un ensemble de catégories. Les concepts sont obtenus par une correspondance de Galois avec généralisation par intervalles, ce qui permet de traiter les données de différents types dans un cadre commun. Une mesure de la généralité d'un concept est alors calculée sous une forme commune pour les différents types de variables. Un exemple illustre la méthode proposée.

Summary

This paper deals with hierarchical or pyramidal conceptual clustering methods, where each formed cluster corresponds to a concept, i.e., a pair (extent, intent). We consider data presenting real or interval-valued numerical values, ordered values and/or probability/frequency distributions on a set of categories. Concepts are obtained by a Galois connection with generalisation by intervals, which allows dealing with different variable types on a common framework. In the case of distribution data, the obtained concepts are more homogeneous and more easily interpretable than those obtained by using the maximum and minimum operators previously proposed. A measure of generality of a concept is defined similarly for all these variable types. An example illustrates the proposed method.