

# K-centroids Hierarchical Classes Analysis

Luigi Lombardi  
Department of Cognitive Science  
University of Trento  
lombardi@form.unitn.it

Iven Van Mechelen  
Department of Psychology  
University of Leuven  
Iven.VanMechelen@psy.kuleuven.ac.be

**Keywords:** Two-Way Two-Mode Clustering, Multivariate Analysis.

## Abstract

HICLAS models (De Boeck and Rosenberg, 1988; Van Mechelen, De Boeck, and Rosenberg, 1995; Leenen, Van Mechelen, De Boeck, and Rosenberg, 1999) are structural models for  $N$ -way  $N$ -mode data that do allow for linked and structured classifications of the  $N$  modes. More specifically, in the HICLAS models the equivalence classes of each mode-partition are structurally organized in terms of if-then type partial orders. However, up to now, the hierarchical classes family has been limited to models that imply only an implicit upper bound on the cardinalities of the partitions of the distinct modes by means of a function of the so-called mode-ranks of the models. Yet, researchers sometimes may wish to limit explicitly the classification of the elements of one of the modes to at most  $K$  clusters (whereas partitioning on the other modes is left unconstrained). Two substantive reasons may call for the latter scenario. Firstly, for descriptive purposes a researcher may wish to limit the size of the partition of one mode in order to obtain a simpler hierarchical structure for that mode. Alternatively, in a confirmatory approach, a researcher may have a prior model of the partitioning schema of the selected mode in terms of the number of classes in the partition; in this second case, the main goal is to test the goodness of fit of the constrained model with respect to the data.

In this paper we present a new model for clustering of two-way two-mode binary data. The model, which we call KC-HICLAS (K-Centroids Hierarchical Classes Model), is a novel extension of the hierarchical classes model family. Similar to all the members of this family it includes both a linked and a structured classification of the two modes. However, unlike the existing HICLAS models, KC-HICLAS explicitly limits the partition of the elements of one mode to maximum  $K$  (not necessarily distinct) clusters. We illustrate the new model with two applications to real data sets.

## References

- De Boeck, P., & Rosenberg, S. (1988). Hierarchical classes: models and data analysis. *Psychometrika* **53**, 361–381.
- Leenen, I., Van Mechelen, I. & De Boeck, P. (1999). INDCLAS: A three-way hierarchical classes model. *Psychometrika* **64**, 9–24.
- Van Mechelen, I., De Boeck, P., & Rosenberg, S. (1995). The conjunctive model of hierarchical classes. *Psychometrika* **60**, 505–521.