

Visualization in prediction based on grade correspondence analysis

Olaf Matyja ¹⁾ and Wiesław Szczesny ^{1), 2)}

¹⁾ Institute of Computer Science, Polish Academy of Science, Ordona 21, 01-237 Warsaw, Poland

²⁾ Department of Econometry and Computer Science, Warsaw Agricultural University, Nowoursynowska 166, 02-787 Warsaw, Poland

Abstract:

The grade correspondence analysis (GCA) is used to form a solution of predictive problems with a binary response variable and a vector of nonnegative explanatory variables. At first, the GCA is applied to the matrix of values of explanatory variables corresponding to objects belonging to the training set. Then, this analysis is repeated under the restriction that two non-overlapping blocks are formed, each gathering objects of a common value of the response variable. The results provide a general information on the joint behavior of explanatory variables and their link with the response variable. Moreover, the second analysis produces a GCA based discriminant function, used subsequently to classify new objects or objects belonging to a test set. Each phase of this process is clearly visualized by means of suitable over-representation maps. All stages of predicting are illustrated in the paper by simulated university drop-out problem.

Key words:

classification, correspondence analysis, explanatory variable, grade parameters, latent structure, response variable, stochastic dependence, visual intelligent information system.