Investigating food attitudes of Italian children using Projection Pursuit

Cinzia Franceschini[1] **and Nicola Loperfido**[2]

[1]*Sassari University, Sassari, Italy*
[2]*Urbino University, Urbino, Italy*

Abstract

The University of Gastronomic Sciences (Pollenzo, Italy) investigated the attitude of Italian children towards food and its consumption in school canteens. Data were collected from questionnaires administered to 1108 children in 9 primary Italian schools. We first clustered original data by means of model-based clustering and k-means clustering. Then we used principal component analysis to reduce the number of variables before clustering. We obtained the best clustering using k-means on the data projected onto the directions found using projection pursuit, a multivariate statistical technique aimed at finding interesting low-dimensional data projections. By means of projection pursuit we also obtained the most satisfactory variable selection, compared to other statistical methods implemented in R packages. We conclude that the data at hand make a case for projection pursuit when variable selection for clustering is sought.

Keywords

Cluster analysis, Dimension reduction, Multivariate data, Variable selection.

References

[1] Arabie, P. and Hubert, L. (1994). Cluster analysis in marketing research. In: Bagozzi, R. P.(Eds.), *Advanced methods of marketing research*, (pp.160–189). Blackwell, Oxford.

[2] Fop, M. and Murphy, T.B. (2018). Variable selection methods for model-based clustering. *Statistics Surveys. 12*, 18–65.

[3] Neal, M.R. and McNicholas, P.D. (2024). Flexible Variable Selection for Clustering and Classification. *arXiv:2305.16464v2 [stat.ME]*

[4] Sun, J. (2006). *Projection Pursuit. Encyclopedia of Statistical Sciences* (10).

[5] Yoshikazu Terada (2014). Strong Consistency of Reduced K-means Clustering. *Scandinavian Journal of Statistics. 41*, 913 –931