



Application des cartes de Kohonen aux bases de données nutritionnelles

Sonia Fortin, Pascale Rondeau, Sébastien Marque

► **To cite this version:**

Sonia Fortin, Pascale Rondeau, Sébastien Marque. Application des cartes de Kohonen aux bases de données nutritionnelles. 42èmes Journées de Statistique, 2010, Marseille, France, France. 2010. <inria-00494826>

HAL Id: inria-00494826

<https://hal.inria.fr/inria-00494826>

Submitted on 24 Jun 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

APPLICATION DES CARTES DE KOHONEN AUX BASES DE DONNÉES NUTRITIONNELLES

Sonia Fortin, Pascale Rondeau, Sébastien Marque

Danone Research, Centre Daniel Carasso, RD128, Avenue de la Vauve, F-91767 Palaiseau Cedex

Mots-clés : Apprentissage et classification, Biostatistique.

Résumé

Ces dernières années, les modes de consommation alimentaires ont considérablement évolué, aussi bien au sein des sociétés occidentales qu'au sein des pays en développement. Propre à chacun, leur impact direct en Santé est indéniable. La compréhension des pratiques de consommation est donc naturellement au cœur des problématiques nutritionnelles actuelles.

L'utilité d'identification de groupes de consommateurs n'est ainsi plus à démontrer ; mais aujourd'hui aucune classification ne fait consensus. C'est pourquoi il convient d'en proposer de nouvelles.

S'affranchissant de tout a priori, l'approche non-supervisée s'impose à l'approche supervisée de façon évidente. D'autre part, si les méthodes hiérarchiques peuvent être envisagées, les méthodes non-hiérarchiques présentent l'avantage de donner des groupes distincts les uns des autres sans intégrer la notion de formation séquentielle (pas à pas) de ces groupes et laisser supposer que certains soient seulement le résultat de scissions de « sur-groupes ». Au contraire, les consommateurs peuvent être considérés simultanément.

La méthode des cartes de Kohonen, particulièrement adaptée pour les problématiques de grandes dimensions comme les problématiques nutritionnelles, réunit ces deux conditions : elle est non-supervisée et non-hiérarchique. L'algorithme repose en effet sur un apprentissage non supervisé et elle est, de plus, caractérisée par l'association de réponses voisines. On la définit aussi comme une méthode d'auto-organisation.

Appliquée à des bases de données nutritionnelles accessibles, les cartes de Kohonen ont permis de partitionner les consommateurs selon leurs habitudes alimentaires et de fait d'identifier des profils de consommateurs. La classification de Kohonen : un outil précieux dans la compréhension des problématiques nutritionnelles.

KOHONEN CARDS: APPLICATION ON DATABASE OF NUTRITION

Sonia Fortin, Pascale Rondeau, Sébastien Marque

Danone Research, Centre Daniel Carasso, RD128, Avenue de la Vauve, F-91767 Palaiseau Cedex

Keywords: Learning and clustering, Biostatistics.

Abstract

During last decades, behaviours for food consumption have changed considerably, and that is the case in occidental societies as well as in developing countries. Peculiar for each, their impact on the Health is undeniable and the understanding of food practices is of course in the middle of actual problematic of nutrition.

The need to identify groups of consumers is also not to prove. However, today any clustering has yet reached a consensus. To propose new clustering is also essential.

All a priori set aside, unsupervised approaches are obviously more appropriate than supervised approaches. Moreover, if hierarchical methods could be planned, non-hierarchical methods have the advantage to give separate groups without consider the idea of a sequential algorithm (step by step); and without let to think that some groups are just the result from the scission of bigger subgroups. On the contrary, consumers can be considered simultaneously.

The Kohonen cards method, especially adapted for high dimensions problems like questions of nutrition gathers these two conditions: it is unsupervised and also non hierarchical. Indeed, its algorithm consists on an unsupervised learning and in addition, it is characterized by the congregation of neighboured responses. It is also defined as a self-organization method.

Applied to available database of nutrition, Kohonen cards permit to split consumers according to their dietary behaviours; and also to identify profiles of consumers. The clustering with Kohonen cards: a precious tool for the understanding of questions of nutrition.

Bibliographie

- [1] Hulshof K.F.A.M., Valsta L.M., Welten D.C., LoÈwik M.R.H. (2001) - Analytical approaches to food-based dietary guidelines in the European setting - *Public Health Nutrition* - 4(2B) - 667-672
- [2] Kohonen T. (2001) - Self Organizing Maps, *Springer- 3rd edition*.
- [3] Newby P.K., Muller D., Tucker K. L. (2004) - Associations of empirically derived eating patterns with plasma lipid biomarkers: a comparison of factor and cluster analysis methods - *American Society for Clinical Nutrition* - 759-767
- [4] Scali J., Richard A., Gerber M. (2000) Diet profiles in a population sample from Mediterranean southern France, *Public Health Nutrition* - 4(2) - 173-182.