

Visual Data Mining With Self-organizing Maps for “Self-monitoring” Data Analysis

Sociological Methods & Research

1-15

© The Author(s) 2016

Reprints and permission:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0049124116661576

smr.sagepub.com



Elia Oliver^{1,2}, Iván Vallés-Pérez³,
Rosa-María Baños^{1,2}, Ausias Cebolla^{1,2},
Cristina Botella^{2,4}, and Emilio Soria-Olivas³

Abstract

Data collected in psychological studies are mainly characterized by containing a large number of variables (multidimensional data sets). Analyzing multidimensional data can be a difficult task, especially if only classical approaches are used (hypothesis tests, analyses of variance, linear models, etc.). Regarding multidimensional models, visual techniques play an important role because they can show the relationships among variables in a data set. Parallel coordinates and Chernoff faces are good examples of this. This article presents self-organizing maps (SOM), a multivariate visual data mining technique used to provide global visualizations of all the data. This technique is presented as a tutorial with the aim of showing its capabilities, how it works, and how to interpret its results. Specifically, SOM analysis has been applied to analyze the data collected in a study on the efficacy of a cognitive and

¹ Dep Personalidad, Evaluación y Tratamientos Psicológicos, University of Valencia, Valencia, Spain

² CIBER Fisiopatología Obesidad y Nutrición, Madrid, Spain

³ Intelligent Data Analysis Laboratory, University of Valencia, Valencia, Spain

⁴ Dep. de Psicología Básica, Clínica y Psicobiología—Personalidad, Evaluación y Tratamientos Psicológicos, University Jaume I, Castelló de la Plana, Spain

Corresponding Author:

Elia Oliver, University of Valencia, Dep Personalidad, Evaluación y Tratamientos Psicológicos, Valencia, Spain.

Email: elia.oliver.gasch@gmail.com

behavioral treatment (CBT) for childhood obesity. The objective of the CBT was to modify the eating habits and level of physical activity in a sample of children with overweight and obesity. Children were randomized into two treatment conditions: CBT traditional procedure (face-to-face sessions) and CBT supported by a web platform. In order to analyze their progress in the acquisition of healthier habits, self-register techniques were used to record dietary behavior and physical activity. In the traditional CBT condition, children completed the self-register using a paper-and-pencil procedure, while in the web platform condition, participants completed the self-register using an electronic personal digital assistant. Results showed the potential of SOM for analyzing the large amount of data necessary to study the acquisition of new habits in a childhood obesity treatment. Currently, the high prevalence of childhood obesity points to the need to develop strategies to manage a large number of data in order to design procedures adapted to personal characteristics and increase treatment efficacy.

Keywords

data science, visual data mining, self-organizing maps, tutorial, artificial neural networks

Introduction

In psychology, as in other sciences, large amounts of data are analyzed in order to draw conclusions about them. Many methods can be used to analyze these data, such as factorial analysis and logistic regression, which are classic methods, or more complex methods, such as neural networks or support vector Machines, which are nonlinear. In addition, there are the so-called visual techniques, which can provide visual intuition about data. They can be divided into two groups: univariate, which produce representations of one variable (e.g., histograms, box-and-whiskers plots, etc.), and multivariate, which try to show the relationships among several variables (e.g. parallel coordinates, self-organizing maps [SOM], etc.)

Parallel coordinates is a classic technique used for visualizing high-dimensional space (Zhou et al. 2008), and it is quite useful when the data set is small but not when it deals with thousands of observations and dozens of variables. For these cases, the SOM algorithm is preferred due to its high power of synthesis. The SOM representation technique has been used in different fields, such as medicine (Cattinelli et al. 2012; Chang & Teng, 2007; Rosado-Muñoz et al. 2013), computer security (DeLooze 2004;

Pachghare, Kulkarni, and Nikam 2009), engineering (Frey 2012; Kohonen et al. 1996; Panapakidis et al. 2013), and so on.

Although SOM has been widely used in many applications, it is still used very little in psychology. This article briefly discusses the structure of this neural network model, the method for obtaining the parameters, and the way to interpret the graphs obtained. To explain SOM, a “toy” data set will be used that was generated by the authors to make initial contact with this visual tool easier. To reflect the potential use of this tool in psychology, a case in this field is also presented, using data from a study on a childhood obesity treatment. Changing behavior is the target of the most efficient obesity interventions, and for this reason, data related to childhood obesity have been chosen for this article as an example of applying the SOM to analyze “self-monitoring” data.

Childhood obesity is one of the greatest public health problems in the twenty-first century (Branca, Nikogosian, and Lobstein 2007). Dietary and physical activity self-records are referred to as the “cornerstone” of behavioral weight control programs because they are considered useful techniques to collect information on patients’ behaviors and the acquisition of new habits and to evaluate the treatment effects (Baños et al. 2009, 2011; Oliver et al. 2013). Specifically for childhood obesity treatment, dietary self-records usually include the registering of the type and amount of food and beverages, social situations and the place where the intake occurs, and the emotions related to the intake moment. For the physical activity self-register, the most common variables registered are the type of exercise done, the time of the implementation, and the subjective fatigue associated with the activity. People in treatment record these variables several times a day in order to analyze their progress in modifying their habits. Adherence to self-register procedures is an important indicator of successful weight management. A consistent relationship has been found between self-register adherence and success in both losing weight and maintaining weight loss (Baker and Kirschenbaum 1998; Collins, Kashdan, and Gollnish 2003). However, the large amount of data recorded throughout the treatment makes it necessary to use analysis procedures that can take many variables into account. SOM analysis techniques allow a global view of all the data.

Three hypotheses were formulated in the present study.

- SOM techniques will be useful for exploring the data collected through self-record systems in a childhood obesity treatment.
- SOM techniques will be able to analyze the adherence to self-records in a childhood obesity treatment.

- SOM analysis will be able to analyze the dietary intake and level of physical activity in children who completed a treatment focused on modifying their habits to control their weight.

This article is organized as follows: In the Method section, the basic elements of the SOM algorithm are explained, and an example of the use of this technique is presented using a synthetic data set. Next, the SOM algorithm is applied to a real problem (childhood obesity) in order to extract useful information from the data. This section will describe the participants recruited and the measurements used to analyze the treatment effects. The Results section will focus on the presentation and explanation of the data obtained through the application of SOM analysis in the specific example of childhood obesity treatment. Finally, general conclusions will be summarized in the Discussion section.

Method

SOM

SOM is a specific artificial neural network (ANN) whose purpose is to find groups of patterns (clustering) and compress the information from high-dimensional data into geometric relationships on a low-dimensional representation that allows the visualization and correlation of complex patterns. An ANN is a computational model based on human brain functioning. The main characteristic of an ANN is its ability to acquire information from the environment and improve its performance, taking into account a prescribed model that constitutes the learning paradigm (Haykin 2009).

In a SOM algorithm, the neurons (or nodes) are ordered in two layers: the input layer (composed of N neurons, one for each input variable) and the competition layer (also called the second layer, composed of a topological low-dimensional grid of neurons—usually two-dimensional—geometrically ordered). Each input layer neuron is connected with every unit on the competition layer, and, subsequently, an N -dimensional weight vector is assigned to each competition layer unit. Thus, a set of observations is associated with each second layer unit. The algorithm finds the sets that best describe the observation domains. The sets are arranged on the two-dimensional grid, so that similar sets are closer to each other than to different ones (Kohonen et al. 2001).

Two choices have to be made: the map type (hexagonal or rectangular grid, which indicates the neighborhood relation or topology) and the number of neurons (which defines the size of the low-dimensional grid). These

choices depend on the size and dispersion of the input data (Vesanto et al. 1999). Then, a learning algorithm is used to calculate the associated weights for each competition layer neuron. The neuron whose weight vector is closer to the input observation is called the winning neuron or *best-matching unit* (BMU).

Once the map training is finished, the two-dimensional map can be visualized. It is also called the “components plane,” and it provides qualitative information about how the input variables are related to each other for the given data set. Furthermore, a “hits map” can be prepared. It represents the number of times each map unit, or neuron, was the BMU for each input register, so that the distribution of the BMU for a given set of data is represented. This information gives an idea of the number of input observations gathered in each neuron, which makes it possible to compare the strength of each unit on the components plane.

The SOM algorithm provides qualitative information to establish relationships between variables, and this is not possible with other methods. Given the difficulty of analyzing the data set of the present study, SOM was proposed to extract knowledge. Specifically, SOM analysis allows a large amount of data to be summarized in an easily interpretable graphical representation. Moreover, SOM analysis provides an unsupervised modeling approach, meaning that no a priori hypotheses need to be formulated by users at the beginning of the study. These characteristics make it possible to obtain unbiased results, and unanticipated relationships between different variables can freely emerge.

Obtaining useful information in a relatively complex multidimensional database is extremely straightforward when using this model and applying this interpretation methodology. For this reason, SOM analysis is the preferred procedure for our study due to its ability to provide a compact and unbiased representation of multiple data. Indeed, SOMs provided greater insight about the acquisition of new habits, and they highlighted unpredicted existing relationships.

Analyzed Problem: The Childhood Obesity Treatment

The problem to be addressed is the analysis of data from adolescent and preadolescent children suffering from overweight and obesity, who were receiving a weight loss treatment focused on developing healthy eating habits and increasing their daily physical activity. There were two treatment conditions with the same components: (a) treatment supported by a web platform and (b) traditional treatment (only face-to-face; for a more detailed

description, see Baños et al. 2009, 2011, 2013; Oliver et al. 2013). The only difference between the two conditions is the use of the web platform. The final goal in both conditions was to increase the consumption of healthy foods, reduce the consumption of unhealthy foods, and increase the level of physical activity in order to gain control over their excess weight and promote healthier habits. Moreover, we expected that children in the condition supported by the web platform would show greater results in modifying their habits than children receiving the traditional face-to-face treatment.

Participants

The total sample was composed of 47 children (32 boys) ranging from 8 to 13 years old ($X: 10.48$; $SD: 1.56$), 25 participants in the traditional treatment and 22 participants in the treatment supported by the web platform. The sample was recruited from a Child and Adolescent Cardiovascular Risk Unit. No significant differences were observed in the groups based on sex and age. The body mass index mean was 28.8 ($SD = 3.59$). Regarding weight, z -scores adjusted for sex and age were calculated, with $z = 2.75$ ($SD = 0.29$).

Variables and Measurements

Dietary and physical activity self-registers were used. Dietary self-registers included the type and amount of food and beverage, the social situation and place where the intake occurs, and the emotions related to the intake moment. Physical activity self-registers included the type of exercise, the time of implementation, and the subjective fatigue associated with the activity. Children in traditional treatment compiled the self-register weekly using a traditional “paper-and-pencil” procedure, while children in the treatment supported by the web platform compiled the self-register using an electronic personal digital assistant (PDA; for a more detailed description, see Oliver et al. 2013). Self-register variables were the same in both procedures.

The study included nine variables to be analyzed: treatment condition, adherence to the self-register, fruit intake, vegetable intake, carbohydrate intake, protein intake, dairy intake, fast food intake, and physical activity.

The first variable, “condition,” refers to the treatment condition to which participant is randomly assigned. The variable “adherence” indicates the amount of eating information and physical activities that children include in the self-register during the complete treatment. Treatment was 10 weeks long. During this period, children were asked to introduce information about

eating (breakfast, lunch, dinner, and snacks consumed) and the physical activities that they carried out in their daily lives. Therefore, this variable ranged from zero (not register) to infinity, depending on the number of foods consumed and registered and the activities practiced.

In order to analyze the progress in the food treatment, information about different nutritional groups was registered: fruit consumption (“FS” variable); vegetable consumption (“VS” variable); carbohydrate consumption (“HCS” variable; e.g., bread, cereal, and pasta); consumption of meat, fish, legumes, and eggs (“Prot” variable); dairy consumption (“LS” variables, e.g., milk, cheese, and yoghurt); and consumption of fat foods (“Gr” variable, e.g.; chocolates, cakes, sweets, fast food, chips, sweet beverages, etc.). Physical activity (“AF” variable) was also analyzed as the number of activities the children carried out in their daily lives (e.g., walk, run, play football, dancing, skating, basket, jumping, etc.).

Results

SOM was implemented using a free Matlab Toolbox that is available in <http://www.cis.hut.fi/projects/somtoolbox/download/>

To train the SOM, several parameters were varied:

1. *Initialization*: Two types of initializations were used: (a) random: The initial weights follow a normal distribution (with a zero mean and variance equal to one) and (b) based on principal components analysis (Kohonen 1982).
2. *Training algorithm*: Two types of algorithms were used: (a) *batch*: The adjustment values of the different weights are accumulated across all the training items and (b) *online training*: Weights and bias values are adjusted for each training item.

To obtain the best SOM, two error measures were used: (a) topographic error, measures the ability of SOM to hold neighborly relations between original patterns and the SOM projections and (b) quantization error, gives an idea of the SOM’s accuracy in modeling the data. Among the different versions of the trained SOM, the best version was the one with the minimum product quantization error—topographic error.

SOM analysis was carried out, and the results are summarized in this section in order to find out the distribution of participants with regard to the dietary and physical activity self-registers. Figure 1 shows the distribution of the participants over the whole map. Figure 2 shows the distribution of the

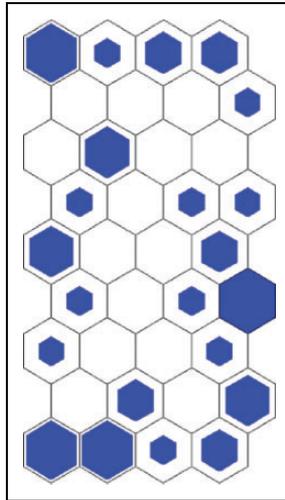


Figure 1. Winners map; this map shows the distribution of the participants along the whole map. This map shows projections corresponding to the different input features used to train the self-organizing maps.

participants according to the variables analyzed: two conditions (traditional vs. treatment supported by a web platform), the adherence to the self-records, intake consumption (fruits, vegetables, carbohydrates, protein, dairy, and fats), and physical activities implemented. The conclusions obtained by observing these figures are summarized below.

- The variable “adherence” is more common in participants who perform fewer “activities” in their daily lives.
- The variables “FS” and “LS” seem to be highly correlated; that is, participants who usually eat more fruit also eat more milk, cheese, and yogurt.
- The variables “HS” and “VS” are highly correlated; that is, participants who usually eat more carbohydrates are likely to eat vegetables.
- Regarding “fat foods,” participants in the traditional condition show high variability, as some of them usually eat many fat foods, whereas some of them do not. However, participants in the treatment supported by a web platform condition do not usually eat fat foods.
- The variable “Prot” does not seem to be correlated with any other variable.

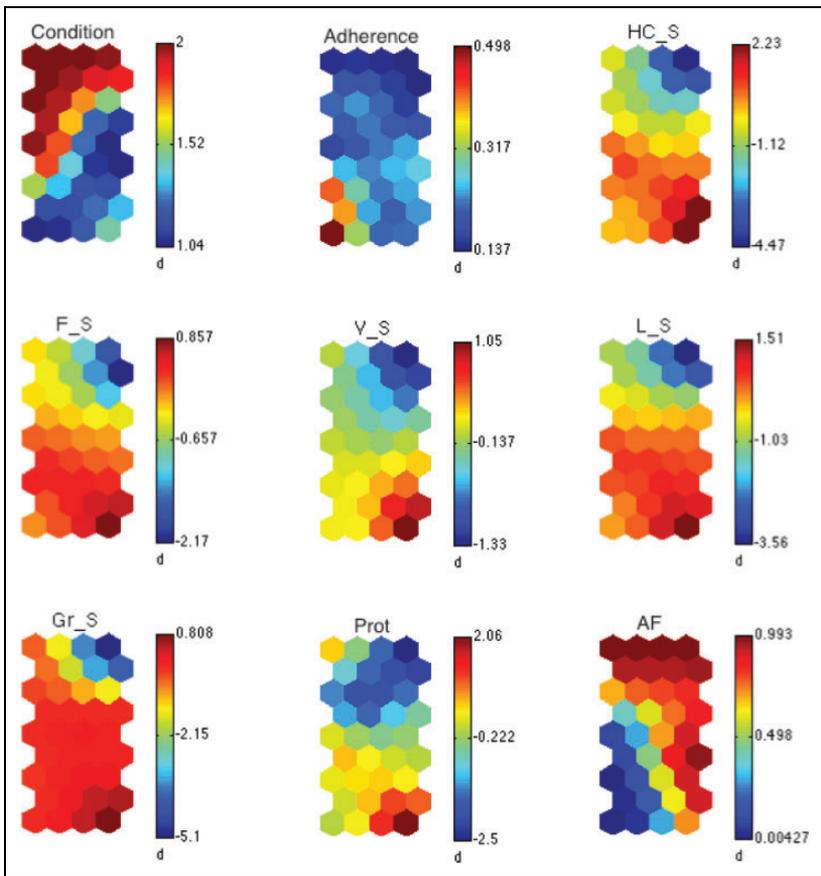


Figure 2. Components map obtained with self-organizing maps algorithm for the synthetic problem data. The colored area inside each hexagon is proportional to the number of input patterns that are more similar to this neuron.

Regarding “adherence,” the values obtained confirm participants’ low adherence to self-register tasks, and this low adherence is even lower for participants in the web platform condition. One possible explanation has to do with the technical problems with the PDA system. Qualitative information provided by children and parents indicated technical difficulties, such as rapid battery discharge, problems with the login and password, or the difficulty of introducing dietary and physical activity information using an electronic pencil. These problems could impede the implementation of electronic

self-registers and shows the importance of making improvements in the devices and software in order to increase participants' adherence.

Discussion

Studies about health behaviors and life styles usually manage a large number of different variables (physical, emotional, cognitive, sociocultural, etc.), and they have to extract multiple data for their analysis. Thus, it is necessary to have efficient strategies to analyze and interpret these data from a multi-variate perspective (Astel et al. 2010). The use of visual data mining emerges as a technique that is able to provide high-quality information about the results obtained and the possibility of designing and developing strategies adapted to personal and clinical characteristics (Wickramasinghe et al. 2011). Visual data mining, and particularly the SOM technique, is a new procedure that takes into account all the variables registered during the treatment and considers these data globally in the analysis procedures. SOM analysis has been used in several medical studies. For example, Wickramasinghe et al. (2011) and Astel et al. (2010) showed that SOM techniques were useful for identifying patterns in patients with diabetes, and these results increased the possibility of developing specific strategies to manage these patients efficiently. SOM analyses have also been used to evaluate the efficacy of a specific screening for assessing the presence of infection in people (e.g., Sun et al. 2011), showing better results than linear discriminant analysis. Other studies have shown the utility of SOM in analyzing the satisfaction of patients in nursing science (Voutilainen et al. 2014), identifying groups of children with different risk profiles for growth development (Schilithz et al. 2014), or obtaining a deeper understanding of ventricular fibrillation (Rosado-Muñoz et al. 2013). However, to the best of our knowledge, no previous study has focused on the use of SOM analysis in childhood obesity treatment.

The final objective of the present study was to obtain a visual representation, in a simple and comprehensive manner, of the relationship between intake and physical activity variables during a treatment focused on modifying habits. Taking into account the hypotheses formulated, SOM techniques were viable and efficient procedures for exploring the data from self-records and analyzing the adherence to these procedures. Moreover, SOM techniques facilitated the understanding of the relationship between the specific treatment implemented (traditional vs. web platform) and its effects on the acquisition of new healthier habits. In this sense, the results showed that children who completed the web platform condition treatment increased their

consumption of healthy foods and reduced their intake of unhealthy food (e.g., fat foods). This information is relevant in designing specific therapeutic strategies to modify lifestyle behaviors.

One of the main limitations of the present study is the low adherence to the self-record procedures, specifically when PDA systems were used. As mentioned above, technical problems can explain this low adherence, such as rapid battery discharge and the difficulty of introducing information using an electronic pencil. The objective of the electronic self-record techniques is to offer a comfortable procedure where specific information can be introduced in an easy and noninvasive way. The final objective is to facilitate the participants' arduous task of introducing the information needed to analyze the treatment progress and make therapeutic decisions. Therefore, our data suggest the importance of developing more ergonomic and easier to use electronic procedures for self-recording behaviors, adapted to the children's lifestyle.

Despite the limitations, this study shows that the SOM is a technique with great potential in this field. Interventions on lifestyle always involve the daily monitoring of many variables because it is assumed that their analysis provides very useful and enriching information. However, until now, much of this information could not be analyzed in a simple way using traditional statistical methods. The SOM provides an easy way to extract information from a lot of data, and we think it will be very useful for analyzing self-monitoring data in psychological interventions.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded in part by the Spanish Ministry of Education, Culture and Sport, Projects ACTIOBE (PSI2011-25767), Excellence in Research Program PROMETEO II (Generalitat Valenciana. Conselleria de Educaci3n, 2013/003), and CIBER Fisiopatolog3a de la Obesidad y la Nutrici3n (ISC III CB06 03/0052).

References

- Astel, A., M. Nikolov, V. Christov, P. Simeonova, and V. Simeonov. 2010. "Chemo-metric Assessment of Clinical Data for Diabetes Mellitus 2 Type Patients Using Self-organizing Maps." *Journal of Environmental Science and Health* 45:560-68.

- Baker, R. and D. Kirschenbaum. 1998. "Weight Control During the Holidays: Highly Consistent Self-monitoring as a Potentially Useful Coping Mechanism." *Health Psychology* 17:367-70.
- Baños, R. M., A. Cebolla, C. Botella, A. García-Palacios, E. Oliver, I. Zaragoza, and M. Alcaniz. 2011. "Improving Childhood Obesity Treatment Using New Technologies: The ETIOBE System." *Clinical Practice & Epidemiology in Mental Health* 7:62-66.
- Baños, R. M., A. Cebolla, C. Botella, E. Oliver, and A. Garcia-Palacios. 2011. "An e-Health system for treatment of childhood obesity: The Etiobe Platform". E-health Communities and Online self-help groups: Applications and usage. Asa Smedberg. ISBN 9781609608668.
- Baños, R. M., A. Cebolla, E. Oliver, C. Botella, and M. Alcañiz. 2013. "Efficacy and Acceptability of an Internet Platform to Improve the Learning of Nutritional Knowledge in Children: The ETIOBE Mates." *Health Education Research* 28: 234-48.
- Baños, R. M., A. Cebolla, I. Zaragoza, C. Botella, and M. Alcañiz. 2009. "Electronic PDA Dietary and Physical Activity Registers in a Weight Loss Treatment Program for Children: A Description of the ETIOBE Personal Digital Assistant System." *Journal of Cybertherapy and Rehabilitation* 2:235-41.
- Branca, F., H. Nikogosian, and T. Lobstein. 2007. *The Challenge of Obesity in the WHO European Region and the Strategies for Response: Summary*. Copenhagen, Denmark: World Health Organization, Regional Office for Europe Copenhagen.
- Cattinelli, I., E. Bolzoni, C. Barbieri, F. Mari, J. D. Martin-Guerrero, E. Soria-Olivas, J. M. Martinez-Martinez, J. Gomez-Sanchis, C. Amato, A. Stopper, and E. Gatti. 2012. "Use of Self-organizing Maps for Balanced Scorecard Analysis to Monitor the Performance of Dialysis Clinic Chains." *Health Care Management Science* 15:79-90.
- Chang, P. L. and W. G. Teng. 2007. "Exploiting the Self-organizing Map for Medical Image Segmentation." Pp. 281-88 in *Computer-based Medical Systems, 2007. CBMS '07. Twentieth IEEE International Symposium, Maribor*.
- Collins, B. L., T. B. Kashdan, and G. Gollnisch. 2003. "The Feasibility of Using Cellular Phones to Collect Ecological Momentary Assessment Data." *Experimental and Clinical Psychopharmacology* 11:73-78.
- DeLooze, L. (2004). "Classification of Computer Attacks Using a Self-organizing Map." Pp. 365-69 in *Proceedings of the 2004 IEEE Workshop on Information Assurance United States Military Academy, West Point, New York*.
- Frey, C. 2012. "Monitoring of Complex Industrial Processes Based on Self-organizing Maps and Watershed Transformations." Pp. 1041-46 in *2012 IEEE International Conference on Industrial Technology (ICIT), Athens*.

- Haykin, S. 2009. *Neural Networks and Learning Machines*. Upper Saddle River, NJ: Pearson Education.
- Kohonen, T. 1982. "Self-organized Formation of Topologically Correct Feature Maps." *Biological Cybernetics* 43:59-69.
- Kohonen, T., E. Oja, O. Simula, A. Visa, and J. Kangas. 1996. "Engineering Applications of the Self-organizing Map." *Proceedings of the IEEE* 84:1358-84.
- Kohonen, T., M. R. Schroeder, and T. S. Huang. 2001. *Self-organizing Maps*. 3rd ed. New York, NY: Springer-Verlag.
- Oliver, E., R. M. Baños, A. Cebolla, E. Lurbe, J. Álvarez-Pitti, and C. Botella. 2013. "An Electronic System (PDA) to Record Dietary and Physical Activity in Obese Adolescents; Data about Efficiency and Feasibility." *Nutrición Hospitalaria* 28: 1860-66.
- Pachghare, V. K., P. Kulkarni, and D. Nikam. 2009. "Intrusion Detection System Using Self Organizing Maps." Pp. 1-5 in IAMA 2009, *International Conference on Intelligent Agent Multi-agent Systems, Chennai*.
- Panapakidis, I., T. Papadopoulos, G. Christoforidis, and G. Papagiannis. 2013. "Analysis of the Electricity Demand Patterns of a Building in a University Campus." Pp. 382-87 in *12th International Conference on Environment and Electrical Engineering (EEEIC), Wroclaw*.
- Rosado-Muñoz, A., J. M. Martínez-Martínez, P. Escandell-Montero, and E. Soria-Olivas. 2013. "Visual Data Mining with Self-organizing Maps for Ventricular Fibrillation Analysis." *Computer Methods and Programs in Biomedicine* 111: 269-79.
- Schilithz, A. O., P. L. Kale, S. G. Gama, and F. F. Nobre. 2014. "Risk Groups in Children under Six Months of Age Using Self-organizing Maps." *Computer Methods and Programs in Biomedicine* 115:1-10.
- Sun, G., S. Abe, O. Takei, Y. Hakozaiki, and T. Matsui. 2011. "A Novel Non-contact Infection Screening System Based on Self-organizing Map with K-means Clustering." *Database Theory and Application, Bio-Science and Bio-Technology* 258: 125-32.
- Vesanto, J., E. Alhoniemi, J. Himberg, J. Kiviluoto, and J. Parviainen. 1999. "Self-organizing Map for Data Mining in Matlab: The Som Toolbox." *Simulation News Europe* 25:54.
- Voutilainen, A., T. Kvist, P. R. Sherwood, and K. Vehviläinen-Julkunen. 2014. "A New Loot at Patient Satisfaction: Learning from Self-organizing Maps." *Nursing Research and Practice* 63:333-45.
- Wickramasinghe, K., D. Alahakoon, P. Schattner, and M. Georgeff. 2011. "Self-organizing Maps for Translating Health Care Knowledge: A Case Study in Diabetes Management." Pp. 162-71 in *Advances in Artificial Intelligence. Volume 7106 of the series Lecture Notes in Computer Science*. Springer.

Zhou, H., X. Yuan, H. Qu, W. Cui, and B. Chen. 2008. "Visual clustering in parallel coordinates. In *Eurographics/IEEE-VGTC Symposium on Visualization 2008, Eindhoven*.

Author Biography

Elia Oliver is a researcher personal at the University of Valencia (Department of Personality, Evaluation and Psychological Treatment) and a member of Labpsitec research group. She finished her doctoral PhD in 2015 focused on the implementation and testing the efficacy of a cognitive and behavioral program for obesity children supported by ICTs. Furthermore, she is working on the use of ICT's for promoting healthy eating behavior and physical activity in children and adolescents.

Iván Vallés is an electronics of telecommunications engineer at University of Valencia. Vocation for data analysis, especially deep learning and ensemble modeling. Currently working in McKinsey & Company.

Rosa-María Baños is a full professor in psychopathology at the University of Valencia (Department of Personality, Evaluation and Psychological Treatment). She finished her doctoral PhD in psychology from the University of Valencia in 1987 and specialist psychologist clinical psychology since 2004. She is leader of research team Labpsitec-Valencia (www.labpsitec.es). Her research interests include psychopathology, the treatment of psychological disorders, and the application of Information and Communication Technologies (ICTs; virtual reality, augmented reality the Internet and mobile apps) to psychology. She is also director of the master "Interventions for Eating Disorders and Personality Disorders" at University of Valencia. She has published over 100 articles and has been principal investigator in more than 20 research projects.

Ausiàs Cebolla is an assistant professor at University of Valencia (Department of Personality, Evaluation and Psychological Treatment) and he is a member of Labpsitec research group. His work is focused on two research lines, the first is about evidence-based mindfulness and compassion interventions for different disorders, and how this skills are related with several well-being indicators. Another research is about healthy lifestyles and its relation with obesity and inactivity. He has coauthored two books about mindfulness (*Mindfulness y ciencia*) and compassion (*La ciencia de la compasión*). Furthermore, he has collaborated in several impact publications and congress. Coeditor of the international journal *Mindfulness and Compassion*.

Emilio Soria-Olivas received a PhD degree from the Universitat de València, Valencia, Spain, in 1997. He is a full professor at the University of València. His research is centered in the application of Advanced Machine Learning Methods to improve decision taking in different scenarios: economic, social, health, and so on. Currently

he is working on deep learning (representation learning) and automatic machine learning (machine learning without parameters).

Cristina Botella is a full professor of clinical psychology at Jaume I University (UJI) (Spain). Director of Labpsitec (www.labpsitec.es). Head of one groups CIBER (Centre for Network Biomedical Research) in the area of Biomedicine and Health Sciences. She is also the director of the doctorate programme in psychology at UJI. She has been principal investigator in more than 40 research projects and published over 150 articles. Her main line of research is the treatment of psychological disorders and the use of ICTs (virtual reality, augmented reality the Internet and mobile apps) to promote health and well-being. She was the president elect of the International Association of CyberPsychology, Training & Rehabilitation in 2009. She received the PhD award at Valencia University in 1983. The Annual CyberTherapy Excellence in Research Award at Cybertherapy 2008. The “Social Board Award for Excellence in Research” at University Jaume I in 2009. First prize “Scientific Dissemination” of Banco Santander 2011; Prize “William Evans Visiting Fellow,” Otago University in New Zealand 2012; First Prize LLGA Cities Pilot the Future in San Francisco 2013 (USA).