

Chapter 32

Visual Data Mining in Physiotherapy Using Self- Organizing Maps: A New Approximation to the Data Analysis

Yasser Alakhdar

University of Valencia, Spain

José M. Martínez-Martínez

University of Valencia, Spain

Josep Guimerà-Tomás

University of Valencia, Spain

Pablo Escandell-Montero

University of Valencia, Spain

Josep Benitez

University of Valencia, Spain

Emilio Soria-Olivas

University of Valencia, Spain

ABSTRACT

The basis of all clinical science developments is the analysis of the data obtained from a particular problem. In recent decades, however, the capacity of computers to process data has been increasing exponentially, which has created the possibility of applying more powerful methods of data analysis. Among these methods, the multidimensional visual data mining methods are outstanding. These methods show all the variables of one particular problem on the whole allowing to the clinical specialist to extract his own conclusions. In this chapter, a neural approximation to this kind of data mining is shown by means of the valuation analysis of the knee in athletes in the pre- and post-surgery of the anterior cruciate ligament, studying variables of force and measurements at different distances of the knee.

INTRODUCTION

Clinical data provide information that enables us to establish new and better diagnoses and treatments for certain pathologies. In physical therapy the analysis of clinical data is of particular

significance because of its wide range of research options in relation to patients, pathologies and their treatment, and the important number of variables that can influence the evolution of an injury and its recovery. A complete and accurate analysis of the data can contribute to the development of more effective therapies and treatments for the patient. It should be noted that data in the clini-

DOI: 10.4018/978-1-4666-2455-9.ch032

cal area (and specifically in physiotherapy) have a set of special characteristics compared to other kinds of data (DeMets et al., 2006):

1. The human body and its interaction with its environment is one of the most complex systems that exist. Therefore, it is logical to consider these relationships as non-linear, that is, an increase of one cause does not lead to a proportional increase of its effects.
2. There are many variables that define the evolution of an injury. We understand that the more the problem is simplified the greater the number of errors that are contained in model. Consequently, there is a non-linear problem owing to a large number of uncorrelated variables.
3. The patients' data collection sheets of a particular pathology or disease may be incomplete or contain errors of measurement.
4. The clinical data increase gradually over time, so the best models to apply are those that can take into account new data reliably.

These characteristics entail that the use of classical statistical models (multivariate regression, logistic regression, clustering algorithms) is not the most suitable given that these models do not highlight the subjectivity and the noise that, in many cases, affect these data. An alternative for the knowledge extraction from data is visual data mining. In this case a multidimensional visualization of the variables on the whole is considered (Chun et al., 2008).

In this way, the clinical specialist could extract his own conclusions with no need of learning the underlying of the models that the data specialist develops. As an example, if the results of a logistic regression are exposed, it is necessary to know what is understood by confidence intervals for the parameters, which are the initial hypothesis of the model as well as the interpretation of the model output. A visual approximation to the data analysis avoids all these problems since the clinical

specialist observes the different behaviours that include his data in a direct way.

One of the most powerful multidimensional data visualization tools is the self-organizing map, a kind of neural model, which is described in the next section. Later is shown the application of these neural networks in a physiotherapy problem showing the potential of these methods. Finally the conclusions of the work are exposed.

SELF-ORGANIZING MAPS

The Self-Organizing Map (SOM) is a neural network proposed by Teuvo Kohonen in 1984 (Kohonen, 2000; Haykin, 2008). Neurons are arranged in two layers, Figure 1: an input layer, formed by n neurons (one neuron for each input variable) and an output layer in which information is processed; this second layer is usually arranged in a two-dimensional structure.

Neurons of the output layer are characterised by a weight vector with the same dimension as the input vector. For instance, neuron i,j (i -th row, j -th column) is characterised by the weight vector $w_{ij} = [w_{ij}^1 \ w_{ij}^2 \ \dots \ w_{ij}^n]$. Similar input patterns are mapped close each other in the output layer (Kohonen, 2000). Algorithm procedure can be summarized, as follows (Haykin, 2008; Kohonen, 2000):

1. Weight initialisation.
2. Choice of an input pattern $x = [x_1 \ x_2 \ \dots \ x_n]$.
3. Measurement of the similarity between weights and inputs. If the Euclidean distance is taken into account, then the similarity measure is given by
$$d(w_{ij}, x) = \sum_{k=1}^M (w_{ij}^k - x_k)^2.$$
4. The most similar neuron to the input pattern is called Best Matching Unit (BMU).

6 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/visual-data-mining-physiotherapy-using/73462

Related Content

Data Field for Hierarchical Clustering

Shuliang Wang, Wenyan Gan, Deyi Liand Deren Li (2011). *International Journal of Data Warehousing and Mining* (pp. 43-63).

www.irma-international.org/article/data-field-hierarchical-clustering/58637

Mining and Analysis of the Traffic Information Situation in the South China Sea Based on Satellite AIS Data

Tianyu Pu (2023). *International Journal of Data Warehousing and Mining* (pp. 1-25).

www.irma-international.org/article/mining-and-analysis-of-the-traffic-information-situation-in-the-south-china-sea-based-on-satellite-ais-data/332864

Scientific Data Management and Visualization: A Service-Driven Integration Approach

Mariana Goranova (2016). *Big Data: Concepts, Methodologies, Tools, and Applications* (pp. 2047-2073).

www.irma-international.org/chapter/scientific-data-management-and-visualization/150256

Cotton Leaf Disease Detection by Feature Extraction

Savita N. Ghaiwatand Parul Arora (2016). *Research Advances in the Integration of Big Data and Smart Computing* (pp. 89-104).

www.irma-international.org/chapter/cotton-leaf-disease-detection-by-feature-extraction/139397

Medical Document Clustering Using Ontology-Based Term Similarity Measures

Xiaodan Zhang, Liping Jing, Xiaohua Hu, Michael Ng, Jiali Xia Jiangxiand Xiaohua Zhou (2008). *International Journal of Data Warehousing and Mining* (pp. 62-73).

www.irma-international.org/article/medical-document-clustering-using-ontology/1800