Artificial Neural Networks in Physical Therapy

Pablo Escandell-Montero
University of Valencia, Spain

Yasser Alakhdar
University of Valencia, Spain

Emilio Soria-Olivas
University of Valencia, Spain

Josep Benítez
University of Valencia, Spain

José M. Martínez-Martínez
University of Valencia, Spain

INTRODUCTION

Clinical data contain valuable information that can be used to establish better diagnostics and treatments. Typically, these data have been analysed using classical statistics methods; however, the amount of data collected has rise rapidly in the last years making necessary the use of more powerful analysis tools. Artificial neural networks (ANNs) are a computational tool which can be used for multiple inference tasks in data analysis such as data categorisation, prediction, modelling and visual data mining. These methods, that have been widely applied in fields as diverse as dairy industry (Goyal, 2012), geosciences (Shuangcheng & Du, 2003) or process control (Hussain, 1999), can provide help for improving the treatments and decisions of physical therapy practitioners. However in the field of physical therapy and sport medicine there is nearly total absence of this kind of models.

This work tries to make known these methods so that physical therapy specialists can apply them to improve the conclusions obtained with other techniques. To this end, we present two of the most widely used neural models for data analysis together with practical examples in the physical therapy field. These models are a generalization of classical statistical methods: multilayer perceptron (MLP) that is an extension of logistic regression and self-organizing maps (SOMs) that are a type of ANN used for visualizing low-dimensional views of high-dimensional data, akin to multidimensional scaling. It should be remarked that these methods do not have the limitations of the classical ones (for example, they can model nonlinear relationships in data or display data with more than three variables).

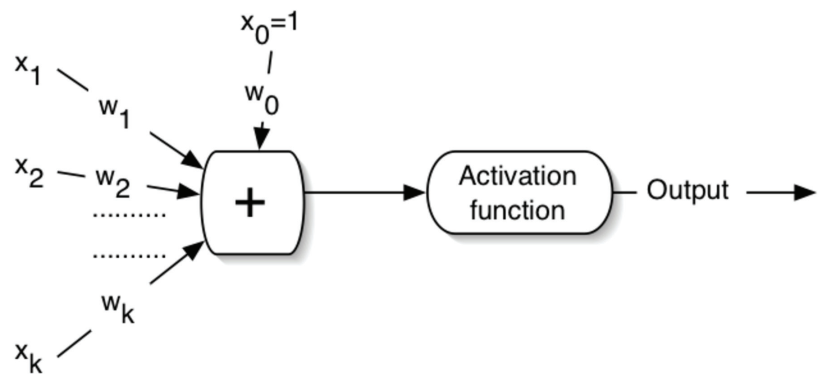
DOI: 10.4018/978-1-4666-5888-2.ch625

BACKGROUND

ANN models take their inspiration from the brain. The brain is an information processing device that has some impressive abilities in many domains, for example: vision, speech recognition, and learning, to name three. Neurons in our brain are parallel connected to other neurons forming a massive parallel computer like machine. ANNs are designed in a way to seek the style of computing of the human brain. As a result, ANNs are powerful enough to solve a variety of problems that are proved to be difficult with conventional methods (Alpaydin, 2004).

In contrast to the conventional mathematical logic, the human thinking process is characterized because it is imprecise, fuzziness, but adaptive. ANNs are designed in a way to mimic most of these characteristics. They learn by examples and exhibit strong adaptation to unseen cases (generalization). Furthermore, informa-

Figure 1. Scheme of a neuron



tion is processed in a massive parallel manner and the neural models are able to deliver a robust behaviour (fault-tolerance) (Arbib, 2003; Bishop, 2007).

MULTILAYER PERCEPTRON

MLP is an artificial neural network formed by elementary processing units, the so-called neurons. A typical neuron model is shown in Figure 1 (Haykin, 2009).

The main components of the model are:

- **Sum function:** It carries out a linear combination of the neuron inputs through the use of a set of coefficients, known as synaptic weights. Being $w=[w_0, w_1, ..., w_k]$ the vector of coefficients, and $x=[1, x_1, ..., x_k]$ the input vector, the sum function is given by the scalar product of both vectors.
- **Activation function:** It is a non-linear function, which gives the network its non-linear nature. The most used activation functions are the sigmoid function (its values ranging between 0 and 1) and the hyperbolic tangent, which ranges between –1 and +1.

When using a neural model, the goal is to find the optimal synaptic weights to solve the problem. The process of searching the optimal weights is known as network learning or training. Most of the algorithms used in training ANNs employ some form of gradient descent. This method basically consists of taking the derivative of a cost function with respect to the network weights and then adjusting those weights in a gradient-related direction. A thorough description of gradient descent learning algorithms is outside of the scope of this article; however, good reference books are Arbib (2003), Bishop (2007), and Haykin (2009). Besides gradient descent, other optimization methods can be used for training ANNs such as evolutionary methods (Ferreira, 2006) and simulated annealing (Da & Xiurun, 2005).

As it can be inferred from Figure 1, a neuron without an activation function is equivalent to a multivariant analysis. Therefore, a non-linear combination should be more powerful than a multivariant analysis (Ripley, 1996). Regarding the network structure, neurons are arranged in layers to form an MLP. The first layer is known as input layer, and the last one is called output layer. All the other layers are called hidden layers (Arbib, 2003). This kind of arrangement enables the neuron outputs to be used as inputs to neurons of following layers (non-recurrent network) and/or previous layers (recurrent networks). Figure 2 shows a typical MLP structure.

This network is used to perform a mapping between two data sets. Some typical problems that can be solved with this model are outlined bellow (Haykin, 2009; Bishop, 2007):

- **Classification:** Starting from the value of variables that define the patient (physical characteristics, injury parameters, medical history), the objective is to assign a patient to a predetermined group. There are many medical areas

Related Content