Analytica Chimica Acta 635 (2009) 121–131

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Review

Artificial neural networks in foodstuff analyses: Trends and perspectives A review

Federico Marini*

Dipartimento di Chimica – Sapienza Università di Roma, P.le Aldo Moro 5, I-00185 Rome, Italy

ARTICLE INFO

Article history:
Received 28 October 2008
Received in revised form 24 December 2008
Accepted 6 January 2009
Available online 10 January 2009

Keywords:
Food analysis
Chemometrics
Artificial neural networks
Food science

ABSTRACT

Artificial neural networks are a family of non-linear computational methods, loosely inspired by the human brain, that have found application in an increasing number of fields of analytical chemistry and specifically of food control. In this review, the main neural network architectures are described and examples of their application to solve food analytical problems are presented, together with some considerations about their uses and misuses.

Contents

1. Introduction

Artificial neural networks (ANNs) are a set of mathematical methods, often encompassed with artificial intelligence, which in some way attempt to mimic the functioning of the human brain [1]. Their introduction in the literature dates back to the late ‘50s, and was the result of the work of scientists by many different fields including neurophysiologists or eminent mathematicians like John von Neumann. Originally they were designed to be a schematic but as accurate as possible model of the activity of the human brain to capture the essential features that make it more powerful than any existing computer (versatility, adaptive response to external stimuli, highly effective pattern recognition ability even in the presence of noisy data, and so on). Under this respect the pioneering works of McCulloch and Pitts [2], Hebb [3], Rosenblatt [4], Widrow and Hoff [5] deserve to be mentioned. However, the initial enthusiasm of contemporary scientists about the topic rapidly faded during the

* Tel.: +39 06 49913856; fax: +39 06 4457050.
E-mail address: fmmonet@homail.com.

0003-2670/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2009.01.009

70s, mainly due to two events. First of all, the practical difficulties in solving many real-world problems. Secondly, the results of the theoretical study published in 1969 by Minsky and Papert [6], who showed that, in the form they were used at the time (called the *perceptron*), neural networks suffered serious limitations, that could not be overcome in a simple way. Therefore, there was a diffused belief that the era of artificial neural network would soon have come to an end when, in 1982, Hopfield published a paper [7] that was destined to stimulate a resurgence in neural network research, as he introduced two key concepts that allowed to overcome all the limitations identified by Minsky and Papert: the non-linearity between the total input received by a neuron and the output it produces and the possibility of feedback coupling of outputs with inputs. This milestone, together with the introduction of the back-propagation (BP) algorithm in 1986 [8–10] has triggered an explosion of interest, together with a change in paradigm: in recent years, there is greater interest in using neural networks as problem-solving algorithms than in developing them as accurate representations of the human nervous system. Accordingly, they are being successfully applied across an extraordinary range of problem domains, in areas as diverse as finance, medical diagnosis, process control, engineering, geology, weather forecasting, physics and, obviously, chemistry.

Among the different factors that can be cited as responsible for neural networks being applied successfully in many areas so different among themselves, two important ones are without any doubt that they are very sophisticated non-linear computational tools capable of modeling extremely complex functions and that they can *learn by example*: the data structure is automatically learnt from representative data by means of opportunely designed *training algorithms*.

As in many other analytical fields, the use of ANNs for data processing has significantly been increasing for the last 20 years, so that examples of application of this technique to almost every aspect of food analysis can be found in the literature. However, compared to other areas, the diffusion of computational models based on neural networks for food analysis is still at a relatively earlier stage of development, so that on one hand many researchers either do not know about the existence of the technique or ignore its potential for solving food control-related problems, while on the other hand one can find in the literature examples of the misuse of ANNs due to an inadequate knowledge of their principles.

Therefore, the aim of this review is to critically discuss the possibility of applying artificial neural networks for food analysis, by presenting a general introduction to the technique, a description of the main typologies of problems encountered in the field and some examples of solution, and considerations about some key issues. Indeed the kinds of data encountered in food science are various and often rather different among themselves, ranging from problems where a limited number of "clean" variables are measured on a suitable number of samples and the underlying model is almost linear or at least mildly non-linear to situations where many variables, possibly noisy or highly correlated are measured on a small number of samples and the functional relation that one wants to model is heavily non-linear. In such situations, a versatile and adaptive technique like artificial neural networks can provide a better modeling, where traditional chemometric techniques fail.

2. A brief survey of ANN concepts

Artificial neural networks [11–13] are a family of mathematical models that share among themselves the characteristics that their main algorithmic features are somewhat inspired to some issues of the functioning of the human brain. However, it should be stressed that while the idea of mimicking and modeling the real neural activity was a primary goal of the first network models, nowadays this aspect is mostly left aside and neural networks are used more as a mathematical than a biological model. Under this respect, for the sake of its chemical applications, a neural network can be thought of as a way of modeling a functional relationship between a set of input and a set of corresponding output variables:

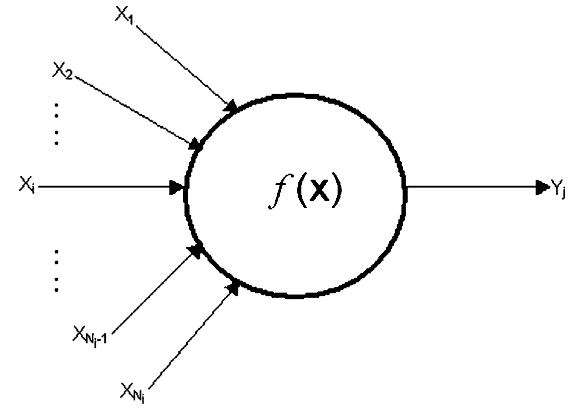

Fig. 1. Schematic representation of an artificial neuron.

$$\mathbf{y} = f(\mathbf{x}) \tag{1}$$

where **x** and **y** are the input and output vectors, respectively, and the symbol *f* means that a functional relationship is sought. Depending on the applications, the **y** vector can represent the sample coordinates on a reduced-dimensionality space (*exploratory analysis*), a binary vectors of class-membership (*classification*) or a real-valued dependent vector (*regression*). All these applications will be described in detail in the following sections.

When looking at the peculiarities of artificial neural networks, what makes them different from traditional mathematical models used in chemometrics is the way the functional relationship described in Eq. (1) is accomplished. Indeed, in neural networks modeling the functional dependence between the output and the input space is described in an implicit way, rather than analytically: the peculiarity of ANNs relies on the fact that they operate using a large number of parallel connected simple arithmetic units (that are called neurons in analogy to their biological equivalent). Mathematically speaking, a neuron can be defined as a non-linear, parameterized, bounded function, so that the variables this function depends on are called the inputs of the neuron and its value is called the output (see Fig. 1). In this framework, parameterization can occur in two different fashions:

1. The parameters are associated to the inputs of the unit, so that a 'global input' of the neuron is built as a linear combination of the inputs x_i, weighted by the parameters (called weights, w_i); the output of the unit is then obtained as a non-linear function (here labeled *f*) of this global input, according to:

$$y = f(w_0 + \sum\nolimits_{i=1}^{n} w_i x_i) \tag{2}$$

2. The parameters are assigned to the neuron non-linearity, i.e. they take part in the definition of the unit. This occurs, for instance, when *f* is a gaussian radial basis function:

$$y = \exp[-\sum\nolimits_{i=1}^{n} \frac{(x_i - w_i)^2}{2w_0^2}] \tag{3}$$

where the $\{w_i\}_{i=1:n}$ are the coordinates of the barycenter of the multivariate normal, and a constant standard deviation, w_0, is assumed along all directions.

As far as now, we have only spoken of individual neurons: the definition of an artificial neural network requires a further passage. Indeed, just as a neuron can be thought as a non-linear function of