Pre issue, CMA4CH 2006, Mediterranean Meeting

Multivariate Analysis and Chemometrics applied to Environment and Cultural Heritage, Nemi (RM), 2-4 October 2006, Italy, Europe

Artificial Neural Networks in Chemometrics: History, Examples and Perspectives

F. Marini, R. Bucci, A.L. Magrì, A.D. Magrì

Chemistry Department, Rome University "La Sapienza"
P.le Aldo Moro 5, I-00185, Rome, Italy

Abstract

Artificial Neural Networks (ANNs) are non-linear computational tools suitable to a great host of practical application due to their flexibility and adaptability. However, their application to the resolution of chemometric problems is relatively recent (early 90s).

In this communication, different artificial neural networks architectures are presented and their application to different kinds of chemometric problems (mainly classification and regression) is discussed by means of examples taken from the authors 'experience, stressing the pros and cons of ANNs with respect to traditional chemometric techniques

Introduction

Artificial Neural Networks (ANNs [1]) are non-linear computational tools designed to reproduce in an abstract and algorithmic way some essential features of the functioning of the human brain, in particular adaptive learning and parallel distributed processing. These two issues, together with the high flexibility resulting from the possibility of representing any non-linear functional mapping, given a suitable choice of the network to be used, would suggest ANNs as a desirable technique to deal with real-world chemometric problems. However, despite the positive characteristics listed above, ANNs still represent a sort of unsolved mystery not only for practitioners, but also for better trained chemometricians.

In this communication, the potential (but also the drawbacks) of using artificial neural networks in some relevant fields of chemometrics (exploratory analysis, classification and regression) will be discussed, by means of examples taken from our own experience. Some of the most used networks architectures will de described and compared and two novel algorithms developed in our lab specifically for class-modeling will be presented.

Methods, i.e. From the artificial neuron to the net

From a computational point of view, artificial neural networks are just one of the possible ways to operate a non-linear mapping between an input and a target space: the peculiarity of ANNs relies on how this mapping is carried out. Indeed, whatever the network chosen, this functional relationship is expressed in an implicit way as the resultant of the parallel elaboration made by several interconnected units (neurons). In particular, different patterns of interconnection give rise to different so called "network architectures", each with its own learning rule.

The most common network architecture is the multilayer feed-forward (see Figure 1), in which neurons are organized in three kind of layers: input, hidden and output: the input units receive information from the outside world, usually in the form of a datafile; the intermediate neurons, contained in one or more hidden layers, allow nonlinearity in the data processing; the output layer is used to provide and answer to a given set of inputs [2]. This architecture can be used both for classification and regression; in principle it has the best generalizing ability but is the most

exposed to the risk of overfitting the data. Moreover, it requires a lot of samples to be trained efficiently.

On the other hand, Kohonen Self Organizing Maps[3] use the computational paradigm of competitive learning to operate a graphical mapping of the data onto a 2D grid of discrete coordinates: this architecture is optimal for exploratory analysis, as a non-linear alternative to PCA (and different than NLM). When a predicted vector is assigned to each of the positions of the 2D-grid, the network is called a counterpropagation network (CP-NN [3]). With respect to other techniques, CP-NN are less prone to overfitting and converge in significantly less iterations; however, their generalizing ability is limited by the size of the 2D grid.

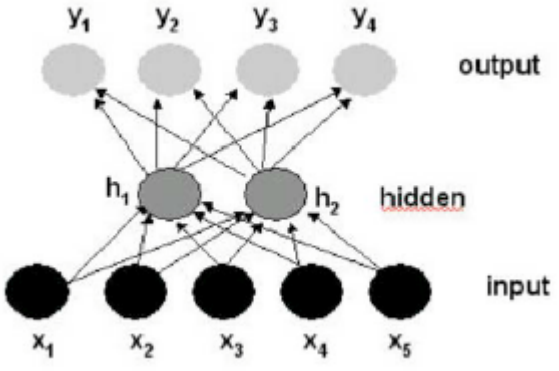

Fig. 1; Scheme of a multilayer feed-forward artificial neural network.

Results and perspectives

It is apparent from the previous paragraph that the choice of a correct architecture can be crucial for the overall performance of the computational tool. In particular, we have demonstrated the potential of a counterpropagation architecture for problems involving classification (wheat [4], rice [5]) while, on the other hand, feed-forward architecture significantly outperformed the others when applied to regression problems (e.g. in QSAR [6]). In any of these cases, neural-based tools significantly outperformed traditional chemometric techniques, confirming the potential of this technique to deal with real-world problems, potential that we're also trying to increase by designing novel algorithms capable of widening the range of topics ANNs can cope with.

.

References

1) J. Zupan, J. Gasteiger, Neural Networks in Chemistry and Drug Design, 2nd ed., Wiley VCH, Weinheim, 1999, ISBN: 3527297790

2) D.E. Rumelheart, G.E. Hinton, R.J. Williams, Learning Internal Representations by Error Back-propagation. In D.E. Rumelheart and J.L. McClelland (eds), Parallel Distributed Processing. Explorations in the microstructure of cognition, MIT Press, Cambridge, 1986, p.318-362, ISBN: 026268053X.

3) J.Zupan, M. Novic, I. Ruisánchez, Kohonen and counterpropagation artificial neural networks in analytical chemistry, *Chemom. Intell. Lab. Syst.*, 38 (1997) 1-23

4) F. Marini, R. Bucci, A.L. Magrì, A.D. Magrì, R. Acquistucci, R. Francisci, Classification of 6 durum wheat cultivars from Sicily using Artificial Neural Networks, submitted to *Chemom. Intell. Lab. Syst.*

5) F. Marini, J. Zupan, A.L. Magrì, On the use of counterpropagation artificial neural networks to characterize Italian rice varieties, *Anal. Chim. Acta*, 510 (2004), 231-240

6) F. Marini, A. Roncaglioni, M. Novic, Variable selection and interpretation in structure-affinity correlation modelling of estrogen receptor binders, *J. Chem. Inf. Model.*, 45 (2005), 1507-1519