Chapter 3
Neural Network Models:
Usefulness and Limitations

Eliano Pessa
University of Pavia, Italy

ABSTRACT

The Artificial Neural Network (ANN) models gained a wide popularity owing to a number of claimed advantages such as biological plausibility, tolerance with respect to errors or noise in the input data, learning ability allowing an adaptability to environmental constraints. Notwithstanding the fact that most of these advantages are not typical only of ANNs, engineers, psychologists and neuroscientists made an extended use of ANN models in a large number of scientific investigations. In most cases, however, these models have been introduced in order to provide optimization tools more useful than the ones commonly used by traditional Optimization Theory. Unfortunately, just the successful performance of ANN models in optimization tasks produced a widespread neglect of the true – and important – objectives pursued by the first promoters of these models. These objectives can be shortly summarized by the manifesto of connectionist psychology, stating that mental processes are nothing but macroscopic phenomena, emergent from the cooperative interaction of a large number of microscopic knowledge units. This statement – wholly in line with the goal of statistical mechanics – can be readily extended to other processes, beyond the mental ones, including social, economic, and, in general, organizational ones. Therefore this chapter has been designed in order to answer a number of related questions, such as: are the ANN models able to grant for the occurrence of this sort of emergence? How can the occurrence of this emergence be empirically detected? How can the emergence produced by ANN models be controlled? In which sense the ANN emergence could offer a new paradigm for the explanation of macroscopic phenomena? Answering these questions induces to focus the chapter on less popular ANNs, such as the recurrent ones, while neglecting more popular models, such as perceptrons, and on less used units, such as spiking neurons, rather than on McCulloch-Pitts neurons. Moreover, the chapter must mention a number of strategies of emergence detection, useful for researchers performing computer simulations of ANN behaviours. Among these strategies it is possible to quote the reduction of ANN models to continuous models, such as the neural field models or the neural mass models, the recourse to the methods of Network Theory and the employment of techniques borrowed by Statistical Physics, like the one based on the Renormalization Group. Of course, owing to space (and mathematical expertise) requirements, most mathematical details of the proposed arguments are neglected, and, to gain more information, the reader is deferred to the quoted literature.

DOI: 10.4018/978-1-4666-9770-6.ch003

INTRODUCTION

As well known, for a long time ANNs have been successfully employed for building models and analyzing data in the hard sciences and engineering. Instead, within the context of Economics and Management Sciences the use of ANN-based tools is more recent and often the traditional statistical analysis tools, such as linear regression, are still widely employed, and preferred to ANNs. In any case, the growing interest in ANNs, showed in these domains, has mainly been due to their evidenced advantages as *statistical analysis* tools, rather than to their usefulness as *modeling* tools (see, for instance, DeTienne, DeTienne and Joshi, 2003; Palocsay and White, 2004; Scarborough and Somers, 2006; Somers and Casal, 2009; Minbashian, Bright and Bird, 2010). Namely, the most quoted reasons supporting the interest in ANNs include:

1. The ability to account for nonlinear relationships;
2. The absence of the need for a specification of the model underlying the data under consideration;
3. The ability to work in absence of suitable statistical assumptions about the error terms (which, on the contrary, must hold in order that a linear regression can make sense);
4. The possibility of adapting the structure and operation of ANNs as a function of the data to be dealt with;
5. The absence of the need for a specification of data statistical distribution;
6. The ability to deal with data in which there are important correlations between different .variable values.

The claimed superiority of ANNs as statistical analysis tools, when compared mostly with linear regression, has been supported by a number of papers, some of which dealing with business data (see, e.g., Razi and Athappilly, 2005; Abu Bakar and Tahir, 2009; Moradzadehfard, Motlagh and Fathi, 2011), but many others with non-business data (we list here only few references, such as Prybutok, Yi and Mitchell, 2000; Li, Wunsch, O'Hair and Giesselmann, 2001; Shi, Lee, Lee, Ho, Sun, Wang, 2012). However, despite the good performance sometimes evidenced by ANNs on data analysis, most people keep a doubtful attitude towards them, essentially related to the fact that they still appear as *non-transparent* tools. Namely, the interpretation, and mainly the explanation in terms of cause-effect relationships, of their answers is difficult. And, notwithstanding the consistent advances in the symbolic interpretation of ANNs operation (see Sun and Alexandre, 1997; Wermter and Sun, 2000; Browne and Sun, 2001; Duch, Setiono and Żurada, 2004; Kamruzzaman and Monirul Islam, 2005; Setiono, Pan, Hsieh and Azcarraga, 2005; Odajima, Hayashi, Gong and Setiono, 2008; Setiono, Baesens and Mues, 2011; Sun, 2013), most researchers still prefer to resort to already known statistical tools, such as regression, advantaged also by the recent availability of new books and better software dealing with nonlinear regression analysis (see on this topic Brown, 2001; Motulsky and Christopoulos, 2004; Bates and Watts, 2007).

As a matter of fact, ANNs first gained popularity as modeling tools in Neuroscience. And this is obvious: in this domain model's biological plausibility is an essential requirement, strictly related to the need for connecting microscopic aspects of biological neuron operation with the macroscopic features observed by psychologists and cognitive scientists. The understanding of this connection is perhaps the main goal included in connectionist's manifesto (Rumelhart, McClelland, 1986; but see also Gluck and Rumelhart, 1990: Ramsey, Stich and Rumelhart, 1991), and the ANNs seem the best (maybe unique) way for reaching this goal. However this goal, when translated into connectionist language, is tantamount to