Chapter 8
Neural Networks

Rosaria Silipo
International Computer Science Institute, Berkeley, USA

8.1. Introduction

There is a wide variety of quantitative methods available for data analysis. A popular approach is based on the simulation of biological nervous systems' functions.

The basic premise of this approach - known as Artificial Neural Networks (ANNs) - is that biological systems perform extraordinarily complex computations without recourse to explicit quantitative operations. Organisms are capable of learning a task gradually over time. This learning property is thought to reflect the ability of large ensembles of neurons to learn through exposure to external stimuli and to generalize across related instances of the signal. Such properties of biological neural systems make them attractive as a model for computational methods designed to process complex data in a flexible manner, that is relatively independent of the task defined.

Two additional advantages of biological nervous systems are the relative speed with which they compute and their robust (i.e., fault tolerant) nature in the face of environmental or internal degradation. Thus, "damage" to a local portion of an ANN usually has little impact on the operation of the neural network as a whole, due to the inherent redundancy of the computational architecture. And often the condition of the incoming signals is less than pristine, corrupted by various sources of environmental degradation such as background noise or missing data.

ANNs represent a family of models, rather than a single technique. Each form of ANN is optimized for a specific set of conditions, analogous to the functional specificity associated with different regions of the brain. This chapter focuses on the most popular ANN paradigms used for quantitative data analysis. A more biological approach to ANNs can be found in [261, 253].

Our discussion begins with an historical overview of ANNs, starting with the initial attempts to computationally simulate biological systems in the 1940's and their evolution into more highly structured ANNs, such as the Multilayer Perceptron, in the 1960's.

We will then proceed to discuss the Back-Propagation learning algorithm (section 8.3), as well as how ANNs generalize to previously unseen data (section 8.4). Although ANNs trained with Back Propagation can be very effective under many different conditions, it is rare that they yield the sort of insight and understanding desirable in a scientific endeavor. In order to provide a more explanatory framework, ANNs can be supplemented with Radial Basis Function units (section 8.5) or with an auto-organizing learning algorithm (Kohonen maps, section 8.6). ANNs can also be used to reduce the dimensionality (i.e., representational complexity) of the data set, analogous to the application of Principal Components Analysis (section 8.7). The chapter concludes with a discussion of how ANN techniques can be applied to time-series data (section 8.8).

8.2. Fundamentals

8.2.1 The Artificial Neuron of McCulloch and Pitts

The biological neuron (fig. 8.1.a) represents the elementary unit of any biological nervous system. Biological neurons appear to be organized in structures, where, after an adequate learning period, they cooperate together to solve a high number of complex tasks. The incoming neuronal fibers (the dendrites) receive electrical signals from the connected neurons via biochemical processes through the synapses. Depending on the synapse's chemical nature, each junction can enhance (excitatory synapse) or reduce (inhibitory synapse) the transmitted signal. If the sum of those incoming electrical signals reaches a threshold, an action potential is fired by the cell through the outgoing fiber (the axon) to other, usually a thousand, connected neurons. After firing the neuron has to wait for a time, called refractory period, before it can fire again. Neurons can differ from each other, regarding for example their refractory period, reaction time, synapses nature, and so on. This makes them play a particular role inside the biological neural structure to which they belong.

The first formal definition of an *artificial neuron* was proposed in [369] (see fig. 8.1.b). The input signals $x_1(t), x_2(t), \ldots, x_n(t)$ at time t from other firing neurons are supposed to be binary and are transmitted through the synapses to the cell body. The effects of the synapses are taken into account by appropriately weighting the input signals. The input $a(t)$ to the neuron is then evaluated as the sum of the received input signals $w_1 \cdot x_1(t), \ldots, w_n \cdot x_n(t)$. Finally a step function $f(a)$ (eq. 8.1) sets the output $o(t+1)$ to 0 (inactive neuron) or $+1$ (firing neuron) according to whether the sum of the received input signals $a(t)$ is above or below a given threshold w_0. The function $f(a)$ is called the *activation function* of the neuron.

$$f(a) = \begin{cases} 1 & \text{if} \quad a \geq w_0 \\ 0 & \text{if} \quad a < w_0 \end{cases} \tag{8.1}$$