Artificial Neural Networks Tutorial

N

Crescenzio Gallo
Department of Clinical and Experimental Medicine, University of Foggia, Italy

INTRODUCTION

The purpose of this article is to introduce a powerful class of mathematical models: the artificial neural networks. This is a very general term that includes many different systems and various types of approaches, both from statistics and computer science. Our aim is not to examine them all (it would be a very long discussion), but to understand the basic functionality and the possible implementations of this powerful tool. We initially introduce neural networks, by analogy with the human brain. The analogy is not very detailed, but it serves to introduce the concept of parallel and distributed computing. Then we analyze in detail a widely applied type of artificial neural network: the feed-forward network with error back-propagation algorithm. We illustrate the architecture of the models, the main learning methods and data representation, showing how to build a typical artificial neural network.

BACKGROUND

Artificial neural networks (Bandy, 1997; Haykin, 1999) are information processing structures providing the (often unknown) connection between input and output data (Honkela, Duch, & Girolami, 2011) by artificially simulating the physiological structure and functioning of human brain structures.

The natural neural network, instead, consists of a very large number of nerve cells (about ten billion in humans), said neurons, and linked together in a complex network. The intelligent behavior is the result of extensive interaction between interconnected units. The input of a neuron is composed of the output signals of the neurons connected to it. When the contribution of these inputs exceeds a certain threshold, the neuron - through a suitable transfer function - generates a bioelectric signal, which propagates through the synaptic weights to other neurons.

Significant features of this network, which artificial neural models intend to simulate, are:

- The parallel processing, due to the fact that neurons process simultaneously the information;
- The twofold function of the neuron, that acts simultaneously as memory and signal processor;
- The distributed nature of the data representation, i.e. knowledge is distributed throughout the network, not circumscribed or predetermined;
- The network's ability to learn from experience.

This last but fundamental capacity enables neural networks to self-organize, adapt to new incoming information and extract the input-output connections from known examples that are the basis of their organization. An artificial neural network captures this attitude in an appropriate "learning" stage.

Despite the great success achieved by artificial neural networks, it is however better to remain aware of the limits of this technology due to the necessary reduction of the real system to be examined.

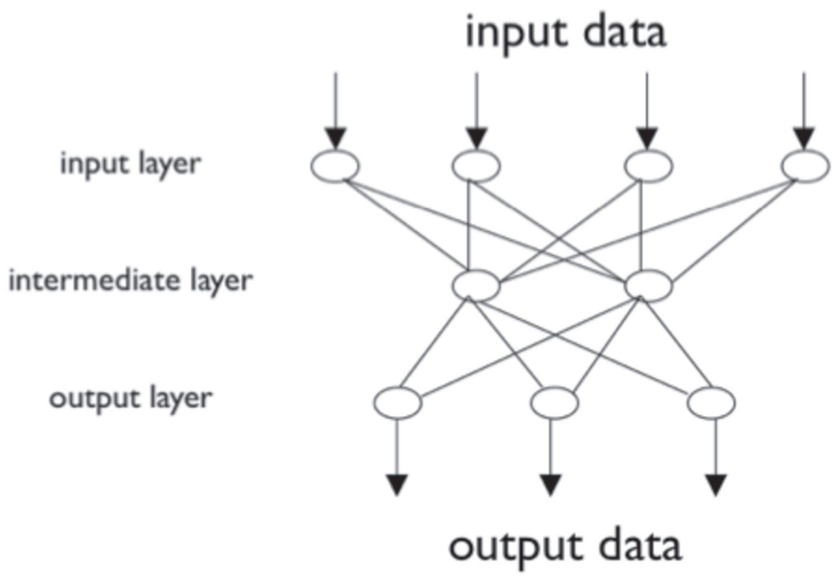
STRUCTURE OF A NEURAL NETWORK

Artificial neural networks are composed of elementary computational units called neurons (McCulloch & Pitts, 1943) combined according to different architectures. For example, they can be arranged in layers (multi-layer network), or they may have a connection topology. Layered networks consist of:

- Input layer, made of n neurons (one for each network input);
- Hidden layer, composed of one or more hidden (or intermediate) layers consisting of m neurons;

DOI: 10.4018/978-1-4666-5888-2.ch626

Figure 1. A basic artificial neuron



- Output layer, consisting of p neurons (one for each network output).

The connection mode allows distinguishing between two types of architectures:

- The feedback architecture, with connections between neurons of the same or previous layer;
- The feedforward architecture (Hornik, Stinchcombe, & White, 1989), without feedback connections (signals go only to the next layer's neurons).

Each neuron (Figure 1) receives n input signals x_i (with connection weights w_i) which sum to an "activation" value y. A suitable transfer (or activation) function F transforms it into the output F(y).

The operational capability of a network (namely, his knowledge) is contained in the connection weights, which assume their values thanks to the training phase.

ARCHITECTURES AND MODELS OF NEURAL NETWORKS

The different configuration possibilities are endless, so the choice of the optimal configuration must be primarily a function of the application target.

Perceptron

The simplest network (Rosenblatt, 1958; Minsky & Papert, 1969) is constituted by a single neuron, with n inputs and a single output. The basic learning algorithm of perceptron analyzes the input configuration (pattern) and — weighting variables through the synapses — decides which output is associated with the configuration.

This type of architecture presents the major limitation of being able to solve only linearly separable problems.

Multi Layer Perceptron (MLP) Networks

The neural network with an input layer, one or more intermediate layers of neurons and an output layer is called Multi Layer Perceptron or MLP (Hornik, Stinchcombe, & White, 1989). This network (see Figure 2) is of feed-forward type and uses, in most cases, the backpropagation learning algorithm. It calculates the weights between layers starting from random values and making small gradual and progressive changes after the network's output errors until the learning algorithm converges to an acceptable error approximation.

There are many other types of more complex structures, but these lie beyond our purposes. Indeed, the feedforward supervised backpropagation architecture

Related Content

A Framework for Understanding Information Technology Resources
Andrew Basden (2008). *Philosophical Frameworks for Understanding Information Systems (pp. 265-308).*
www.irma-international.org/chapter/framework-understanding-information-technology-resources/28085

Swarm Intelligence for Automatic Video Image Contrast Adjustment
RR Aparna (2016). *International Journal of Rough Sets and Data Analysis (pp. 21-37).*
www.irma-international.org/article/swarm-intelligence-for-automatic-video-image-contrast-adjustment/156476

Reasoning on vague ontologies using rough set theory
(). *International Journal of Rough Sets and Data Analysis (pp. 0-0).*
www.irma-international.org/article//288522

Chaotic Map for Securing Digital Content: A Progressive Visual Cryptography Approach
Dhiraj Pandey and U. S. Rawat (2016). *International Journal of Rough Sets and Data Analysis (pp. 20-35).*
www.irma-international.org/article/chaotic-map-for-securing-digital-content/144704

The Use of Geo-Spatial Technology in Handheld Devices for Teaching Geography in a Formal School Context
Pamela Cowan and Ryan Butler (2015). *Encyclopedia of Information Science and Technology, Third Edition (pp. 2636-2646).*
www.irma-international.org/chapter/the-use-of-geo-spatial-technology-in-handheld-devices-for-teaching-geography-in-a-formal-school-context/112680