

ARTIFICIAL NEURAL NETWORKS
MULTILAYER PERCEPTRON

Arkadiusz Gabryś
Matrikel-Nr. 21895081
Department of Computer Science

March 8, 2015

Abstract — Content of this document describes simple and typical multilayer perceptron. Guides from neuron cell to full neural network mathematical model and shows implementation example with both network modeling and training.

Contents

1	Introduction	1
2	Neuron	2
2.1	Cell	2
2.2	Model	3

1 Introduction

Many tasks involving mathematical models and equations cannot be solve by algorithm because of its complexity or limitations in computer calculations. Also such problems exist which are too difficult to be describe by math itself. Sometimes it is even not known if mathematical model exist.

Some of this tasks are trivial for humans and by simulating how the human brain works it is possible to create algorithm that is able to *learn* how to solve a problem. *Learning* mean that mathematical model modify itself in *learning process* to meet certain conditions defined by *training set*. What is *training set* and how *learning process* is done will be covered in next chapters.

Now it is necessary to point that the way brain works is not fully known. But to create simple model our today knowledge is sufficient enough. Structure and mechanism of single neuron is know so based on that mathematical model can be create. Such model is just a generalization of something very complex. Yet good enough to reach our goals and break the bounders of what computer is capable of.

Algorithms uses neural network approach are relativity easy to understand and write in compare to what they can achieve. They are also quite fast. Because of that they are very popular in the filed of machines learning and problem solving.

Popularity of this method has led to creation of many variants and approaches in neural networks field. There exist supervised and unsupervised learning methods, linear and recursive networks. Each approach can be implemented in many different ways.

In this paper the simplest approach with basic *Error Back Propagation* algorithm will be discussed. First the structure and function of single human neuron will be briefly discussed to show how mathematical model is created. After that will be shown how single neurons can be organized together in networks and then in layers. Finally the simplest model of *Multilayer Perceptron* will be discuss along with *Error Back Propagation* learning algorithm. At the end example implementation in *C#* will be presented both with hints, observations and conclusion.

2 Neuron

2.1 Cell



Figure 1: Multipolar neuron cell (*sizes and shapes vary*)

Without going into a details. There exist a few type of neural cell. In our mathematical model we use *Multipolar neural cell* which is shown on Figure 1. This one constitute the majority of neurons in the brain and include motor neurons and interneurons.

The way it works is such that it gather input signals through *dendrites* and produce one output signal through *axon*. The input signals are sum by *nucleus* and only if specific value is obtains the *axon* is activated. Both input and output signal are also modified while they are transmitted to *nucleus* or next cell. Modification depends on *nucleus* and thickness of *neurolemma* (*which is not shown on the figure*). This signal changing part is a place where the *neurons network* knowledge is contained.

2.2 Model

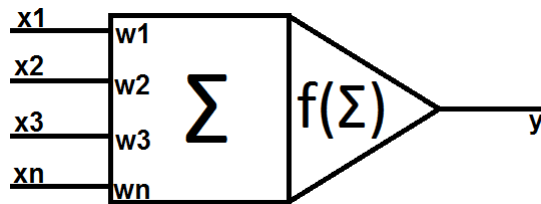


Figure 2: Basic neuron/perceptron model

Such *multipolar neuron cell* mechanism can be easily model. Each *neuron* input is represented by x_n and multiply by corresponding factors w_n which represents change in *neuron* input signal. All signals are sum (that represents *nucleus* function) and then processed by the *activation function* (which is representation of the active or inactive *axon* and also serves other purposes - more about that later in this chapter).

$$f\left(\sum_{i=1}^n x_i w_i\right) = y$$

Equation 1: Basic neuron/perceptron model

Most part of the model is simple and obvious yet the *activation function* need to be discuss in more details.

Activation function may take various form - from simple identity function to step function or more complicated continuous functions. It all depends on the output we want to obtain and/or the problem specification. In today's applications mostly sigmoidal continuous functions are used:

$$y = \text{sgm}(x) = \frac{1}{1 + e^{-\lambda s}}$$

Equation 2: Sigmoidal activation function

Opis wzoru

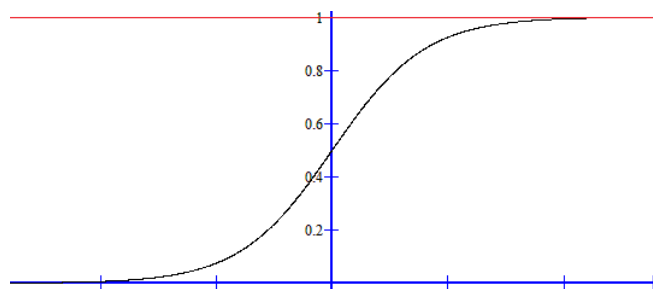


Figure 3: Sigmoidal activation function graph

Opis właściwości wzoru, które można zaobserwować na wykresie