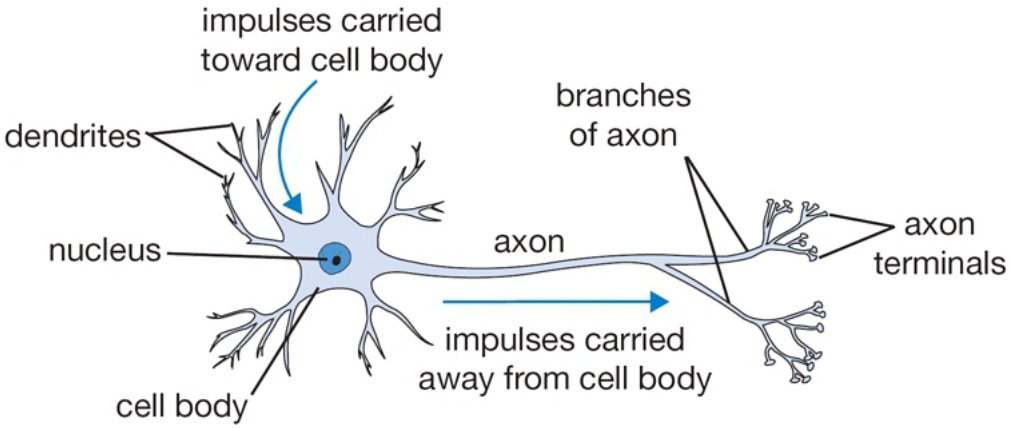
Deep Learning

Paolo Bettelini

Contents

1 Types of neurons

1.1 Brain neurons



1.2 Linear neurons

A linear neuron is very simple and computationally limited in what it can do.

$$y = b + \sum_i x_i w_i$$

The output y is given by the bias b plus the sum of all the input connections x_i multiplied by their weight w_i.

1.3 Binary threshold neurons

Binary threshold neurons output a 1 or a 0 depending on its weighted value.

Given a threshold $\theta = -b$

$$z = b + \sum_i x_i w_i$$

$$y = \begin{cases} 1 \text{ if } z \geq 0 \\ 0 \text{ otherwise} \end{cases}$$

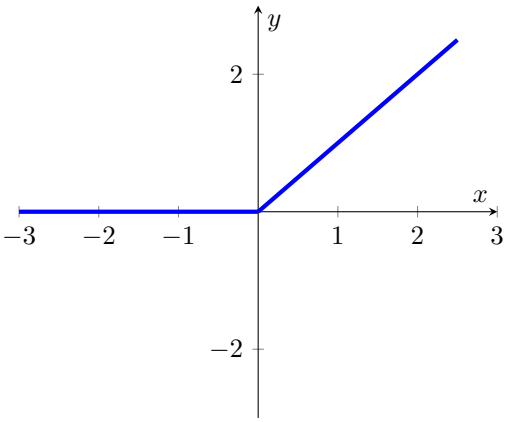
1.4 Rectified Linear Neurons or Linear threshold neurons

They compute a linear weighted sum of their inputs.
The output is a non-linear function of the total input.

Given a threshold $\theta = -b$

$$z = b + \sum_i x_i w_i$$

$$y = \begin{cases} z \text{ if } z > 0 \\ 0 \text{ otherwise} \end{cases}$$



1.5 Sigmoid neurons

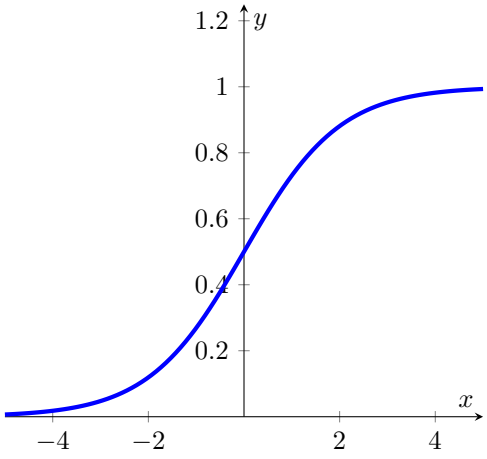
They give a real-valued output that is a smooth and bounded function of their total input.

The logistic function is often used.

Given a threshold $\theta = -b$

$$z = b + \sum_i x_i w_i$$

$$y = \frac{1}{1 + e^{-z}}$$



This function has smooth derivatives that change continuously.
This characteristic makes the learning process easier.

2 Types of learning

2.1 Supervised learning

Each training consists of making the network guess the target output t for a certain input x, given the difference between the correct target and the guess we tweak the network.

There are two type of supervised learning

2.1.1 Regression

The target output is a numeric value of a vector of values.
To describe the error we can compute the square difference between the target output t and the actual output y of the module.
Often the value $\frac{1}{2}(t-y)^2$ is used. The $\frac{1}{2}$ coefficient is there to cancel the 2 out when differentiation is applied.

2.1.2 Classification

The target output is a class or label. Usually either 1 or 0. There could also be multiple labels. [...]

2.2 Reinforcement learning

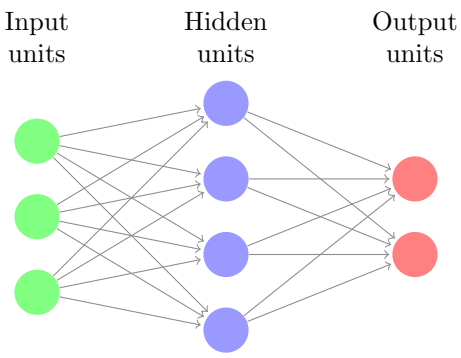
The output is an action of a sequence of actions to maximize the sum of rewards.

2.3 Unsupervised learning

[...]

3 Types of network architecture

3.1 Feed-forward neural network



3.2 Recurrent network

3.3 Symmetrically connected network