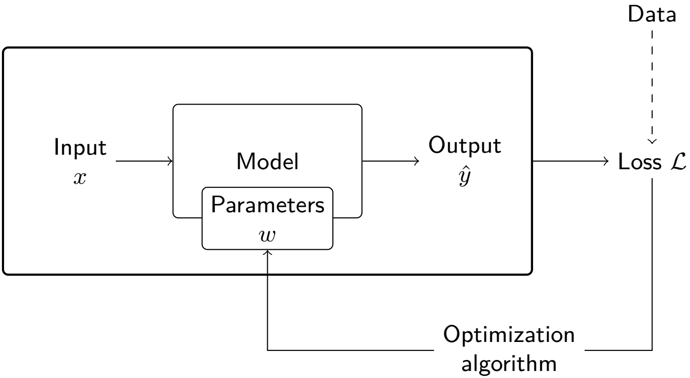
Introduction to deep learning

Leo Bois

August 2022

General Idea



Summary

Toy examples

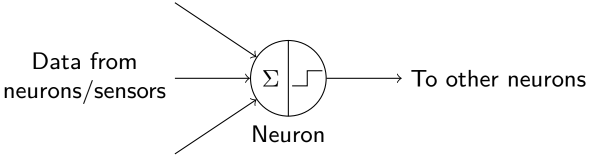
cf Jupyter notebook.

Gradient descent

cf Jupyter notebook.

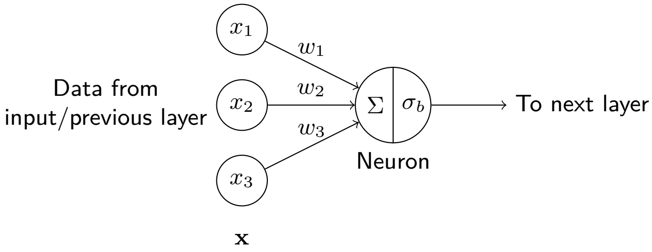
Neural networks

Neural networks are "vaguely inspired by the biological neural networks that constitute animal brains". (Wikipedia)



Single neuron

In machine learning:



Operation implemented:

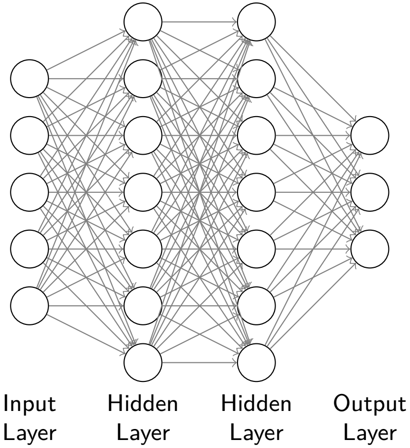
$$\mathbf{x} \mapsto \sigma(\mathbf{w} \cdot \mathbf{x} + b), \quad \mathbf{w} \text{ weights}, b \text{ bias}$$

σ is called the activation function. It is a smooth non-linear function, for instance the sigmoid:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

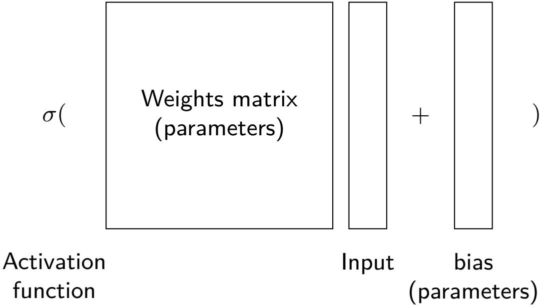
Neural network

A complete neural network looks like this:



Hidden layers

Each hidden layer performs a linear transformation followed by a non-linear activation function:



Such a layer is called *dense* (by opposition to e.g a convolutional layer).
A network with only dense layers is called a multilayer perceptron (MLP).

Backpropagation

For the gradient descent algorithm to work, we have to compute the gradient of the loss with respect to each parameter of the neural network.

As a neural network is a composition of functions, the gradient of the loss with respect to the parameters is computed using the chain rule.
This computation goes from right to left, hence the name *backpropagation*.

In practice, the computation of the gradient is performed by automatic differentiation.

Complete example

cf Jupyter notebook.