NEURAL NETWORKS AND DEEP LEARNING

September 29, 2020

Kerman Sanjuan Malaxechevarria

Index

Chapter 1

Introduction to Neural Networks

1.1 DEFINITION AND USES

A neural network is a network or circuit of neurons, composed by artificial neurons. These networks are used on the field of artificial intelligence, specifically, as a machine learning algorithm. The applications of this algorithm are many, for example, hand-writing recognition, this will be the main example of the first chapter. As we said , neural networks are circuits made out of neurons. But, what does a neuron mean? Formally, an artificial neuron is a mathematical function conceived as a model of biological neurons. Each neuron has it's multiple own inputs and a single output, the output will always depend on the inputs and the "type" of neuron. Now, we will talk about Perceptrons, which are the most simple type of artificial neuron.
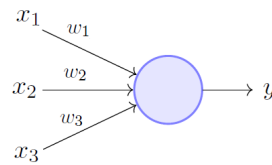
Figure 1.1: Graphical Example of a Neural Network
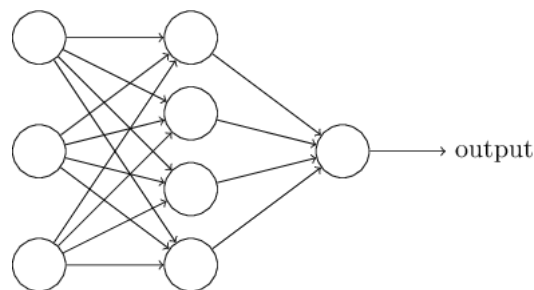
1.2 PERCEPTRONS

The perceptron takes a several binary inputs and produces a single binary output. x_1, x_2, x_3 are the inputs, and w_1, w_2, w_3 the weights , these weights determines the relevance of the input on the output. In other words, a way you can think about a perceptron is that it's

Figure 1.2: Perceptron

Perceptron Model (Minsky-Papert in 1969)

a device that makes decisions by weighing up evidence. Depending on the weights and the threshold, we can get different models of decision making.

$$DesiredOutput = \begin{cases} 0 & \text{if } \sum w_j x_j \leq \text{ threshold} \\ 1 & \text{if } \sum w_j x_j \geq \text{ threshold} \end{cases} \tag{1.1}$$

Obviously, with only one perceptron we are not able to create a neural network, thats because they're structured in such a way that three layers are generated (in most of cases).

Figure 1.3: Basic Structure of a Neural Network
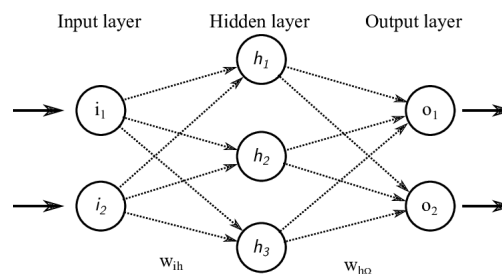
The first layer of perceptrons, also named as input layer, makes simple decisions, wheighing inputs and making simple decisions. Furthermore, the second layer, also called hidden layer, will make more complex decisions. This process will repeat until the last layer, the output one, where the decision will be displayed.

The before mentioned $\sum w_j x_j \leq$ threshold equation is a bit complicated and no intuitive, another way to explain the same is changing $b \equiv -threshold$. The bias can be explained as how easy gets it is to get a 1 as output. For a perceptron with really big bias , it's easy to output a one, instead, with a very large negative number, it will be hard. It's important to mention that, if the Perceptron has two inputs, and both have the same weight, the perceptron will work as NAND gate, taking this into account, we can make every logic gate using perceptrons.

1.3 SIGMOID NEURONS

Sigmoid Neurons are one of the most relevant types of neurons, unlike perceptrons, a little change the weights will just make a little corresponding change on the output. This fact makes sigmoid neurons perfect for learning, because we can adjust the weights according to the desired output. In other words, we could modify the weights and the biases to get the network to behave more in the manner we want. This isn't happens when our network contains perceptrons, in fact, a small change on the inputs will make sometimes to flip completely the output.

The mathematical difference between perceptrons and sigmoid neurons, is the due to the fact that they have different equations. In the case of the sigmoid, the following equation defines the output.

Figure 1.4: Sigmoid Equation

$$\frac{1}{1 + \exp(-\sum_j w_j x_j - b)}.$$

Maybe this isn't the best way to see the difference, so lets see the plot of both neuron types

And now, the plot of the perceptron.

aaa As we see, the reason of the difference of outputs is clear, sigmoid makes a slower change, instead of perceptron which makes a big step.
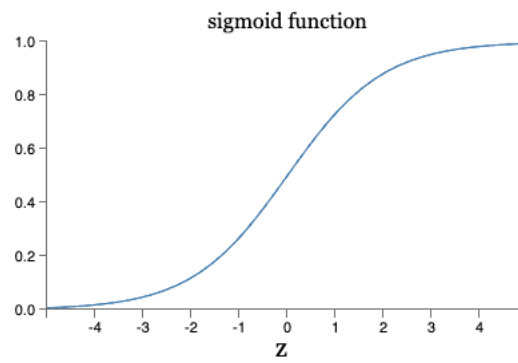
Figure 1.5: Sigmoid Equation Plot

Figure 1.6: Perceptron Equation Plot