

Perceptrons: An Overview

Harjas Monga and Vidhur Kumar

March 30, 2018

"We expect the perceptron to be the embryo of an electronic computer that will be able to walk, talk, see, write, reproduce itself and be conscious of its existence." - The New York Times, 1958.

1 Introduction: What is a Perceptron?

The perceptron is an algorithm for learning a binary classifier: a function that maps its input x (a real-valued **vector**) to an output value $f(x)$ (a single binary value). Shown below is a mathematical definition:

$$f(x) = \begin{cases} 1 & wx + b > 0 \\ 0 & otherwise \end{cases}$$

It's monday morning. You're sitting in your mathematics lecture. The professor asks you to solve the following equation:

$$5 + 2 = ?$$

Within a fraction of a second, you yell "It's seven!" and the professor lauds you. You lean back, thinking "Pfsh! That was easy." You just unknowingly undervalued one of the most incredible tools that you have been blessed with: the human visual system. In each hemisphere of our brain, humans have a primary visual cortex (V1), containing 140 million neurons, with tens of billions of connections between them. There is an entire series of visual cortices doing progressively more complex image processing [1].

We carry in our heads a supercomputer. We do not realize how good we are at making sense of our milieu because all the work is done unconsciously. Only when we attempt to write computer programs to emulate our neural processes is when we truly realize the complexity of the human mind. Nevertheless, the curiosity of scientists persisted and it was in 1958 that Frank Rosenblatt first introduced the idea of a concrete artificial mind. He named it the Perceptron, hoping that "it may eventually be able to learn, make decisions, and translate languages."

It is the idea of the Perceptron that we explore in detail in this paper.

2 Workings: What drives a Perceptron?

2.1 Mathematical Model

A perceptron is nothing more than a function that takes in several inputs and produces an output of either 0 or 1.

The input of perceptron is usually modeled as a column vector. Let $X(n \times 1)$ represent this column vector. The perceptron applies weights to each values in the X and then sums them together. In other words, the perceptron finds a linear combination of the matrix X as show below. At first these weights are random, however, while training the weights get set to the optimal values through a process called back propagation. The inner works of back propagation will be covered later.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}$$

3 Applications: Why are Perceptrons useful?

4 Evolution: What has the idea of a Perceptron led to?