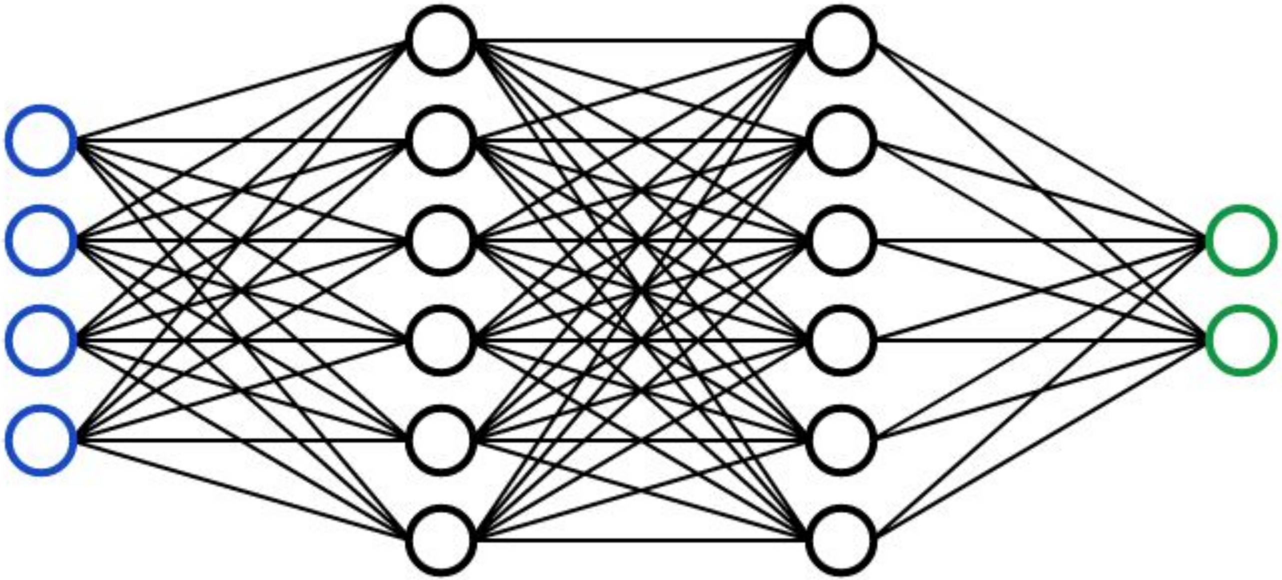
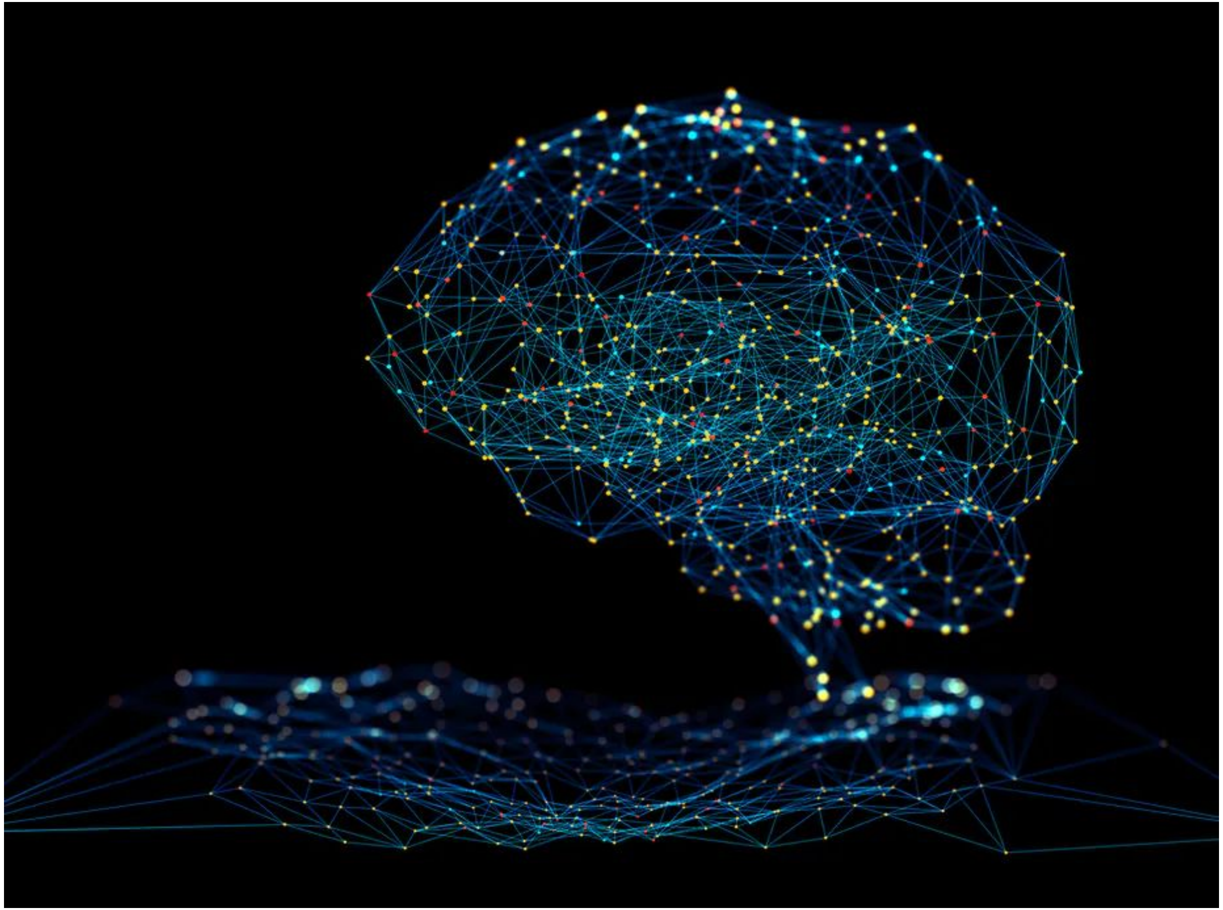
Neural Networks: An Intuitive Introduction

TJ Machine Learning





Neural Network Lecture Series

1. Intuitive Introduction
2. Forward Propagation
3. Backpropagation
    a. Might Include Basic Calculus Lecture before this
4. Hyperparameters
5. Implementation using Tensorflow with Keras

Brief History

- Basic Perceptron Model was developed by Frank Rosenblatt in 1958
- He was a psychologist
- Trying to develop a model for the human brain

FIG. 1 — Organization of a biological brain. (Red areas indicate active cells, responding to the letter X.)

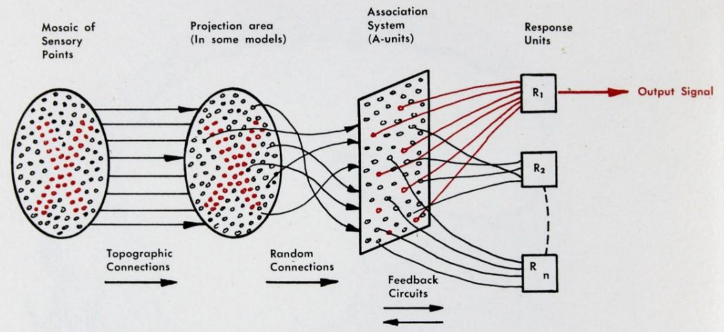
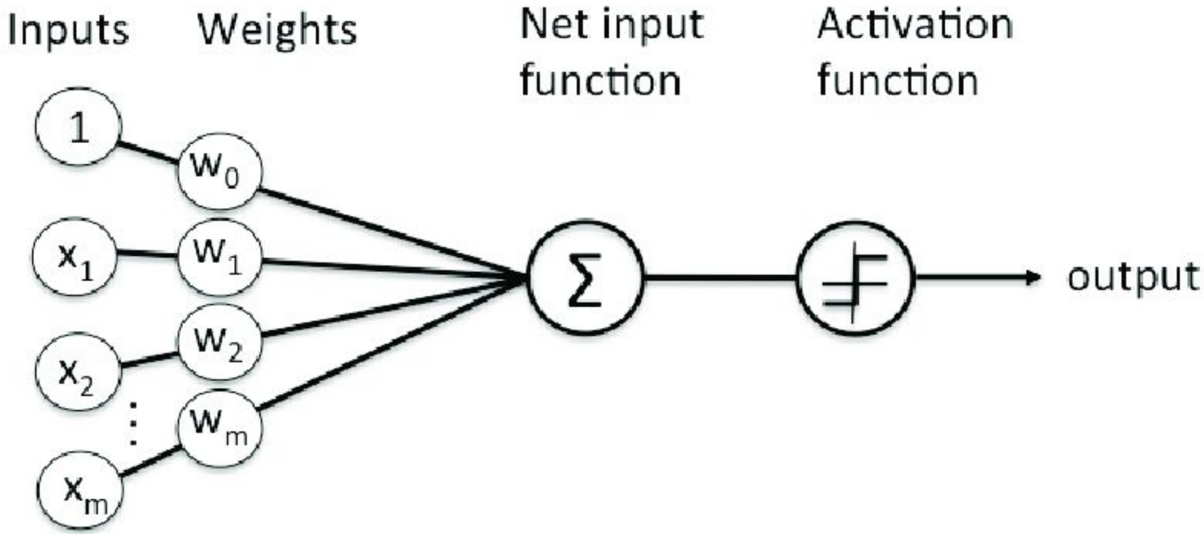


FIG. 2 — Organization of a perceptron.

Perceptron

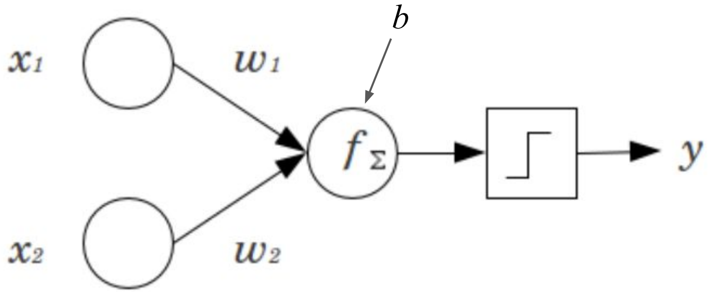
- Simplest Form of a Neural Network



What is a Perceptron?

A perceptron is the most basic unit in a neural network

- Takes in numerical inputs (in the image it takes two, but it can be more)
- Multiplies each input with some weight
- Sums up resulting products and add bias term
- Push resulting sum through activation function (we will use the step function throughout this lecture)



Formalizing the Perceptron

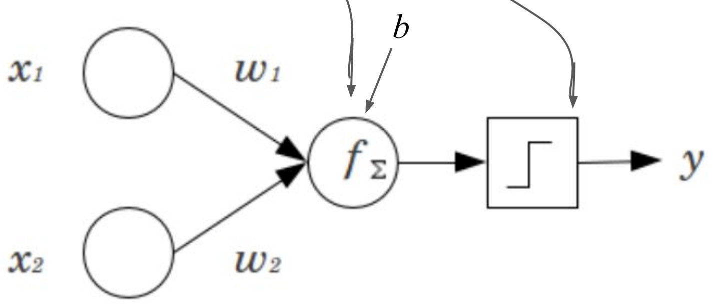
The formal equations that represent the perceptron are:

$$f(x) = w_1 x_1 + w_2 x_2 + b$$

$$y = \begin{cases} 0 & \text{if } f(x) \leq 0 \\ 1 & \text{if } f(x) > 0 \end{cases}$$

Step Function

Note that there would be more terms in f(x) if we had more inputs to the perceptron



Visualization

Perceptrons learn linear functions

- Above the line, they classify data points as 1
- Below the line, they classify data points as 0

The line $x_2 = -x_1 + 1.5$ separates these points well

Rearrange to get $x_1 + x_2 + (-1.5) = 0$

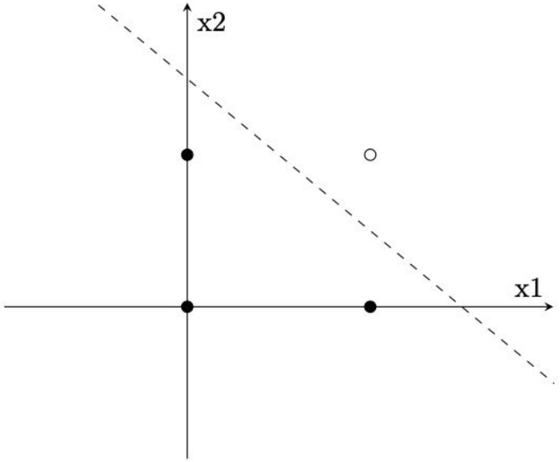
We can then say $\mathbf{f(x) = x_1 + x_2 + (-1.5)}$

For points above line, $f(x) > 0$, so $y = 1$

For points below line, $f(x) < 0$, so $y = 0$

AND Function

x1	x2	y
0	0	0
0	1	0
1	0	0
1	1	1



Learning

where

d - desired output

y - current output

(i_1, i_2) - location of current data point

$$w_1 = w_1 + \alpha(d - y)(i_1)$$

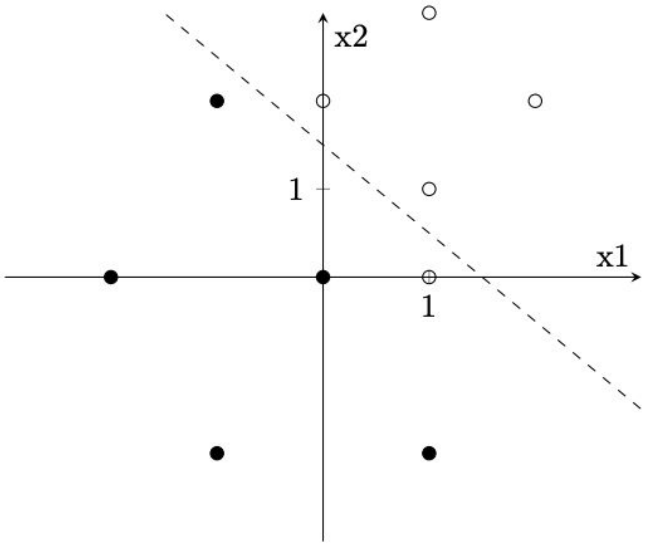
$$w_2 = w_2 + \alpha(d - y)(i_2)$$

$$b = b + \alpha(d - y)$$

Learning

$x_1 + x_2 - 1.5 = 0$

The only misclassified point is at (1,0)



Learning

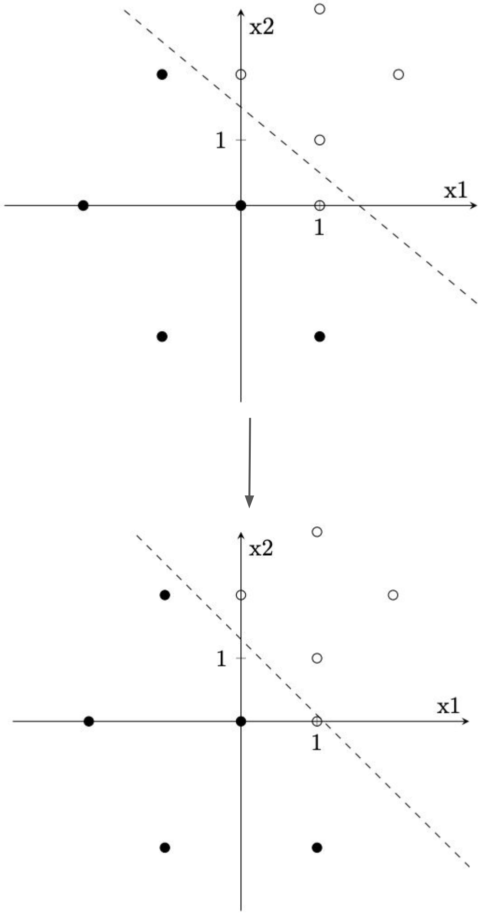
$x_1 + x_2 - 1.5 = 0$

$$w_1 = w_1 + \alpha(d - y)(i_1)$$
$$w_2 = w_2 + \alpha(d - y)(i_2)$$
$$b = b + \alpha(d - y)$$

The only misclassified point is at (1,0)

$$w_1 = 1 + 0.2(1 - 0)(1) = 1.2$$
$$w_2 = 1 + 0.2(1 - 0)(0) = 1$$
$$b = -1.5 + 0.2(1 - 0) = -1.3$$

$1.2x_1 + 1.0x_2 - 1.3 = 0$



Learning

$$w_1 = w_1 + \alpha(d - y)(i_1)$$

$$w_2 = w_2 + \alpha(d - y)(i_2)$$

$$b = b + \alpha(d - y)$$

$1.2x_1 + 1.0x_2 - 1.3 = 0$

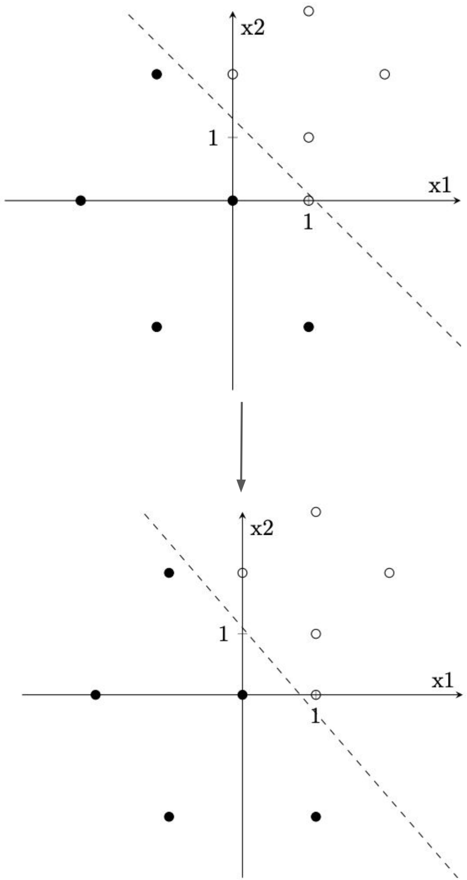
The only misclassified point is still at (1,0)

$$w_1 = 1.2 + 0.2(1 - 0)(1) = 1.4$$

$$w_2 = 1 + 0.2(1 - 0)(0) = 1$$

$$b = -1.3 + 0.2(1 - 0) = -1.1$$

$1.4x_1 + 1.0x_2 - 1.1 = 0$ 🎉 🎉 🎉 🎉



* Note that our process didn't result in the optimal separating line. The process stops as soon as we correctly classify all points.