The Basics of Neural Networks

TM Quest

Overview

What Will we Learn in This Module?

- What are neural networks?
  - The idea behind neural networks.
  - What do we mean when we talk about deep neural networks?
- What are activation functions and weights?
  - How neural networks update their behavior.
  - How activation functions make neural networks interesting.

What are neural networks?

Motivation

Conceptual Description

An (artificial) neural network is based on a collection of connected units or nodes called *artificial neurons*, which loosely model the neurons in a biological brain.

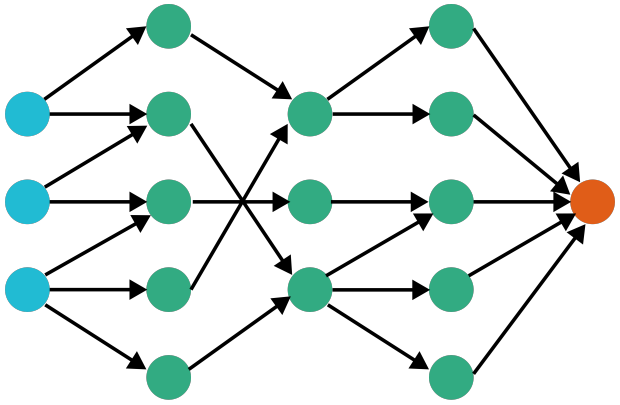
Recent Popularity

Neural nets have become very popular in recent years do to the increase in data volume. They perform well on large amounts of data.

In Practice?

Neural nets are still just machine learning models (for regression and classification) that combine simple pieces into something complicated.

Overview of a Neural Network



Input Layer

Hidden Layers

Output Layer

Weights and activation functions

Weights in Neural Networks

Weighted Input

Each node after the input layer takes its input from the previous layer. It uses weights to put different emphasis on the input based on their importance.

Example

If the input to a node is numbers x and y, then the node create the weighted average

$$w_1 x + w_2 y + b,$$

where w_1, w_2 are weights and b is a bias term. It is the weights w_1, w_2 and the bias term b that are the parameters of the neural network.

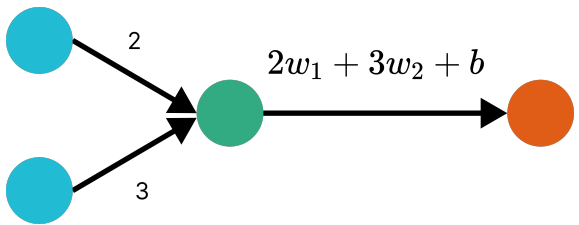
Weights in Neural Networks

Example

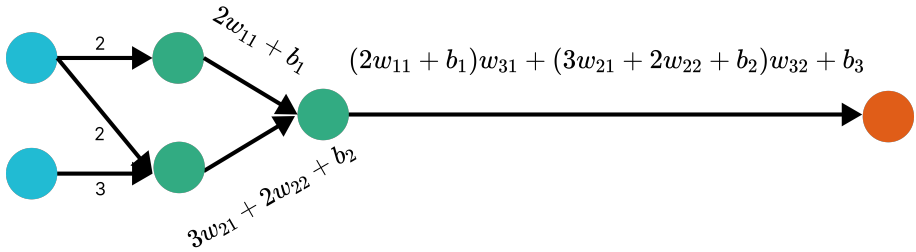
If the input to a node is the numbers x and y, then the node creates the weighted average

$$w_1 x + w_2 y + b,$$

where w_1, w_2 are weights and b is a bias term. It is the weights w_1, w_2 and the bias term b that are the parameters of the neural network.



Why are Weights not Enough?



Summary of the Problem

Problems

- By only using weights (and bias terms) in all the nodes, we are creating linear functions.
- The composition (passing the output of one function into the next one) of linear functions is still linear.
- Deep neural networks for classification in this way can only generate linear decision boundaries.

We need to introduce something that makes the neural nets ”non-linear”!

Activation Functions

Motivation

We introduce activation function between the nodes that introduce non-linearity. Activation functions take the weighted combination from a node and apply a special function to spice things up.

Example

The sigmoid function

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

is an example of an activation function that is used in practice.

In Action

Note!

The activation functions are chosen before the training begins. Hence the weights and bias terms are the only parameter that the model trains on.