Deep Learning

September 14, 2022

Preface

This is my notes about deep learning as I am learning it

Contents

Chapter 1

Introduction to deep learning

1.1 Forward Propagation

Let us build a neural network from scratch and code how it performs predictions using forward propagation. Please note that all deep learning libraries have the entire training and prediction processes implemented, and so in practice you wouldn't really need to build a neural network from scratch. However, this is still good to know to help understanding how they work.
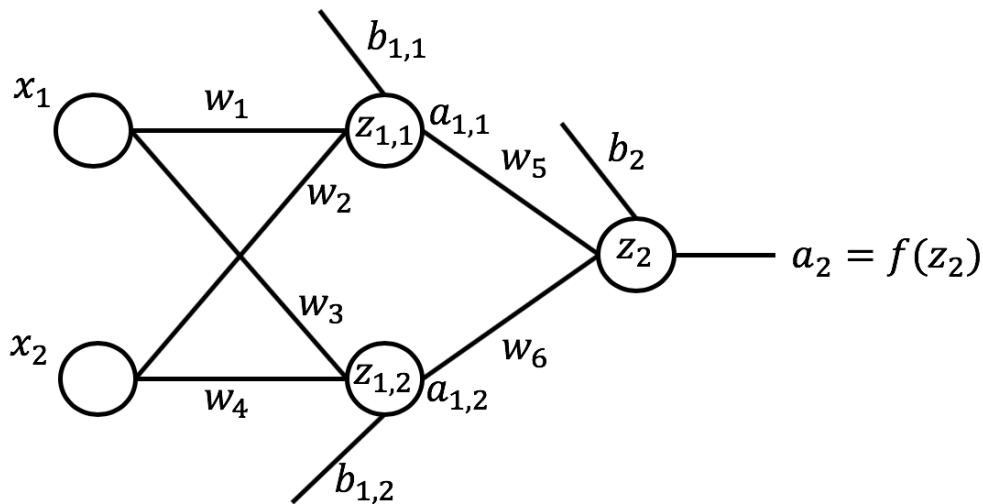

Figure 1.1: A neural network that takes two inputs, has one hidden layer with two nodes, and an output layer with one node.

Let's start by randomly initializing the weights and the biases in the network. We have 6 weights and 3 biases, one for each node in the hidden layer as well

as for each node in the output layer.
We would then calculate the weighted sum of the input at the hidden layer.
It follows that:

$$
\begin{aligned}
z_{1,1} &= x_1 \cdot w_1 + x_2 \cdot w_2 + b_{1,1} \\
z_{1,2} &= x_1 \cdot w_3 + x_2 \cdot w_4 + b_{1,2}
\end{aligned}
\tag{1.1}
$$

Next, assuming a sigmoid activation function, let's compute the activation of the nodes, $a_{1,1}$ and $a_{1,2}$, in the hidden layer.

$$
\begin{aligned}
a_{1,1} &= \frac{1}{1 + e^{-z_{1,1}}} \\
a_{1,2} &= \frac{1}{1 + e^{-z_{1,2}}}
\end{aligned}
\tag{1.2}
$$

These activations would serve as inputs to the next hidden layer or in this case the output layer. Let us compute the weighted sum of these inputs to the node in the output layer.

$$
z_2 = a_{1,1} \cdot w_5 + a_{1,2} \cdot w_6 + b_2
\tag{1.3}
$$

Finally, let's compute the output of the network as the activation of the node in the output layer.

$$
a_2 = \frac{1}{1 + e^{-z_2}}
\tag{1.4}
$$

The activation of the node in the output layer which is equivalent to the prediction made by the network.

Bibliography

[1] M. Artin, *Algebra*. Pearson, 2010.