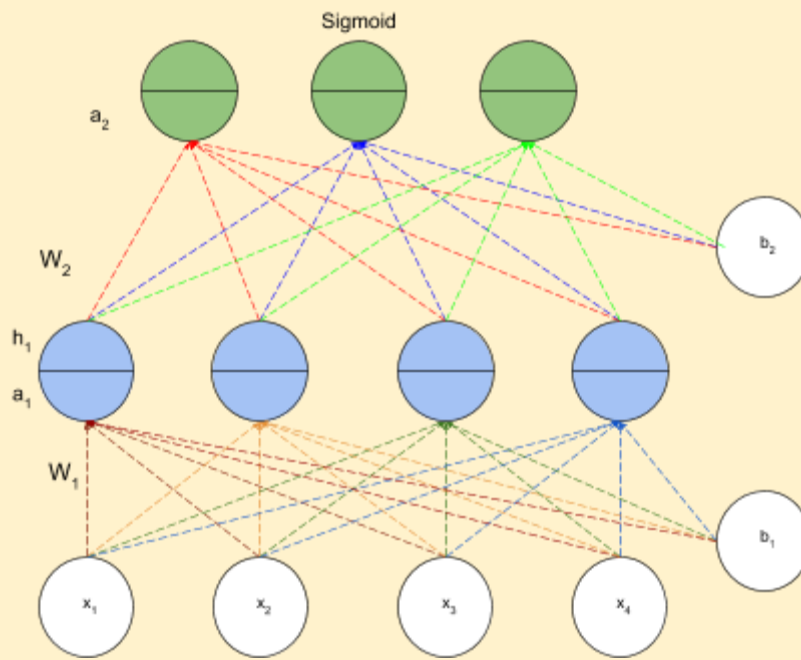


Loss function for Multi-Class Classification

What is the loss function that you can use for a multi-class classification problem

1. Here is an illustration of a sample multi-class classification Neural Network



2. Consider the following values for the parameters

- a. $b = [0 \ 0]$

- b.

$$W_1 = \begin{bmatrix} 0.1 & 0.3 & 0.8 & -0.4 \\ -0.3 & -0.2 & 0.5 & 0.5 \\ -0.3 & 0.1 & 0.5 & 0.4 \\ 0.2 & 0.5 & -0.9 & 0.7 \end{bmatrix}$$

- c.

$$W_2 = \begin{bmatrix} 0.3 & 0.8 & -0.2 & -0.4 \\ 0.5 & -0.2 & -0.3 & 0.5 \\ 0.3 & 0.1 & 0.6 & 0.6 \end{bmatrix}$$

One Fourth Labs

3. Consider a case where $x = [-0.6 \ -0.6 \ 0.2 \ 0.3]$ and true class $y = [0 \ 1 \ 0]$
4. The output values are as follows
 - a. $a_1 = W_1 * x + b_1 = [-0.19 \ -0.16 \ -0.09 \ 0.77]$
 - b. $h_1 = \text{sigmoid}(a_1) = [0.45 \ 0.46 \ 0.49 \ 0.68]$
 - c. $a_2 = W_2 * h_1 + b_2 = [0.13 \ 0.33 \ 0.89]$
 - d. $\hat{y} = \text{softmax}(a_2) = [0.23 \ 0.28 \ 0.49]$
 - e. Cross Entropy Loss
 - i. $L(\Theta) = -\sum_{i=1}^k y_i \log(\hat{y}_i)$
 - ii. $L(\Theta) = -1 * \log(0.28)$
 - iii. $L(\Theta) = 1.2729$
5. Consider another case where $x = [0.6 \ 0.4 \ 0.6 \ 0.1]$ and true class $y = [0 \ 0 \ 1]$
6. The output values are as follows
 - a. $a_1 = W_1 * x + b_1 = [0.62 \ 0.09 \ 0.2 \ -0.15]$
 - b. $h_1 = \text{sigmoid}(a_1) = [0.65 \ 0.52 \ 0.55 \ 0.46]$
 - c. $a_2 = W_2 * h_1 + b_2 = [0.32 \ 0.29 \ 0.85]$
 - d. $\hat{y} = \text{softmax}(a_2) = [0.2718 \ 0.2634 \ 0.4648]$
 - e. Cross Entropy Loss
 - i. $L(\Theta) = -\sum_{i=1}^k y_i \log(\hat{y}_i)$
 - ii. $L(\Theta) = -1 * \log(0.4648)$
 - iii. $L(\Theta) = 0.7661$
7. A quick summary of what we've learned so far
 - a. Given weights, we know how to compute the model's output for a given input
 - b. This is called Forward-propagation.
 - c. Given weights, we know how to compute the model's loss for a given input
 - d. But who will give us the weights?
8. The weights can be obtained from the learning algorithm