Ans. to the Q. No-1

Update rule for regression with backpropagation

Given, we have to train a 2-layer neural network using sigmoid activation function and also with the Mean Square Error (MSE) loss.

The MSE loss function is

$$L = \frac{1}{2}(y - \hat{y})^2$$

The input and output layers -

$$z_1 = x w_1 + b_1$$

$$a_1 = \text{sigmoid}(z_1)$$

$$z_2 = a_1 w_2 + b_2$$

$$a_2 = \text{sigmoid}(z_2) = \hat{y}$$

Updated weights are -

$$w_2' = w_2 - \alpha \frac{\delta L}{\delta w_2}$$

$$w_1' = w_1 - \alpha \frac{\delta L}{\delta w_1}$$

Bias

$$b_2' = b_2 - \alpha \delta b_2$$

$$b_1' = b_1 - \alpha \delta b_1$$

For layer 2 update rule will be as following.

$$\frac{\delta L}{\delta w_2} = \frac{\delta L}{\delta a_2} \cdot \frac{\delta a_2}{(\delta z_2)} \cdot \frac{\delta z_2}{\delta w_2)}$$

$$L = \frac{1}{2} * (y - a_2)^2$$

$$\frac{\delta L}{\delta w_2} = \frac{1}{2} \frac{\delta}{\delta a_2} (y - a_2)^2 \cdot \frac{\delta}{\delta z_2} sig(z_2) \cdot$$

$$\frac{\delta}{\delta w_2} (a_1 w_2 + b_2)$$

$$= \frac{1}{2} 2(y - a_2)(-1) \cdot sig(z_2)$$

$$(1 - sigmoid(z_2)) \cdot a_1$$

$$= -(y-a_2)\ \text{sigmoid}(z_2)\,(1-\text{sigmoid}(z_2)).\,a_1$$

$$\frac{\delta L}{\delta w_2} = -(y-a_2).\ \text{sigmoid}(z_2)\,(1-\text{sigmoid}(z_2)).\,a_1$$

$$\frac{\delta L}{\delta w_1} = \frac{\delta L}{\delta a_2} \cdot \frac{\delta a_2}{\delta z_2} \cdot \frac{\delta z_2}{\delta a_1} \cdot \frac{\delta a_1}{\delta z_1} \cdot \frac{\delta z_1}{\delta w_1}$$

$$= -(y-a_2).\ \text{sigmoid}(z_2)\,(1-\text{sigmoid}(z_2))\,a_1.\,w_2\,(\text{sigmoid}(z_1))(1-\text{sigmoid}(z_1)).\,x$$

As,

$$\frac{\delta z_2}{\delta a_1} = w_2, \quad \frac{\delta a_1}{\delta z_1} = \text{sigmoid}(z_1).\,(1-\text{sigmoid}(z_1))$$

$$\frac{\delta z_1}{\delta w_1} = x$$

$$\therefore \frac{\delta L}{\delta w_1} = -a_1.\,w_2.\,(y-a_2).\,\text{sigmoid}(z_2).\,(1-\text{sigmoid}(z_2)).\,\text{sigmoid}(z_1).\,(1-\text{sigmoid}(z_1)).\,x$$