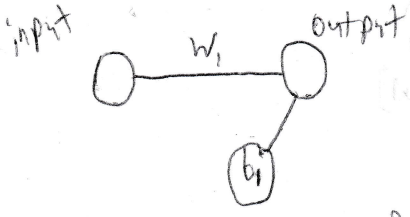
8/11/2024

Neural Net from Scratch Workshop

2 Neurons



① Forward Propogation

$\text{Output} = w_1(\text{input}) + b_1$

② Backward Propogation

$C = (\text{output} - \text{desired\_output})^2$

$\frac{\partial C}{\partial \text{output}} = 2(\text{output} - \text{desired\_output})$

$\frac{\partial \text{output}}{\partial w_1} = \text{input}, \quad \frac{\partial \text{output}}{\partial b_1} = 1$

Chain Rule

$\frac{\partial C}{\partial w_1} = \frac{\partial C}{\partial \text{output}} \frac{\partial \text{output}}{\partial w_1} = (\text{input})\,2(\text{output} - \text{desired\_output})$

$\frac{\partial C}{\partial b_1} = \frac{\partial C}{\partial \text{output}} \frac{\partial \text{output}}{\partial b_1} = 2(\text{output} - \text{desired\_output})$

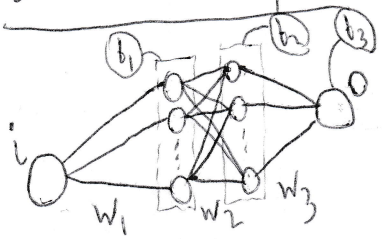
③ Update Weights & Biases

$\alpha = 0.01$ * Typical learning rate

$w_1 = w_1 - \alpha \frac{\partial C}{\partial w_1}$

$b_1 = b_1 - \alpha \frac{\partial C}{\partial b_1}$

2 Hidden Layers



① Forward Propogation

$i\,[1\times 1]$, $w_1\,[h\times 1]$ *h = # Neurons in HL

$b_1\,[h\times 1]$

$z_1 = w_1 .* i + b_1$

$H_1 = \max(0, z_1)$

$w_2\,[h\times h]$, $b_2\,[h\times 1]$

$z_2 = w_2 H_1 + b_2$

$H_2 = \max(0, z_2)$

$w_3\,[h\times 1]$, $b_3\,[1\times 1]$

$z_3 = O = w_3' H_2 + b_3$

② Backward Propogation →

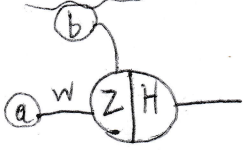
③ Update Weights & Biases

$\alpha = 0.01$

$w_j = w_j - \alpha \frac{\partial C}{\partial w_j}$

$b_j = b_j - \alpha \frac{\partial C}{\partial b_j}$

Single HL Neuron



$z = w(a) + b$

$H = f(z)$ * Activation Function

ReLU Activation Function

$H = \max(0, z)$

Training iters

[150, 500, 1000, 5000]

HL Neurons = 100

$$C = (o - o_{des})^2$$

$$\frac{\partial C}{\partial o} = 2(o - o_{des})$$

$$\frac{\partial o}{\partial W_3} = H_2, \quad \frac{\partial o}{\partial b_3} = 1$$

$$\frac{\partial C}{\partial W_3} = \frac{\partial C}{\partial o}\frac{\partial o}{\partial W_3} = 2(o - o_{des})H_2, \quad \frac{\partial C}{\partial b_3} = 2(o - o_{des})$$

[1×1] [h×1]

$$\frac{\partial H_2}{\partial Z_2} = \begin{cases} 0 & \text{if } H_2(Z_2) = 0 \\ 1 & \text{if } H_2(Z_2) \neq 0 \end{cases}$$

$$\frac{\partial Z_2}{\partial W_2} = H_1, \quad \frac{\partial Z_2}{\partial b_2} = 1$$

$$\frac{\partial C}{\partial W_2} = \frac{\partial C}{\partial o}\frac{\partial o}{\partial H_2}\frac{\partial H_2}{\partial Z_2}\frac{\partial Z_2}{\partial W_2} = \overbrace{2(o - o_{des})(W_3 .* \frac{\partial H_2}{\partial Z_2})}^{[h\times 1]} \overbrace{H_1'}^{[1\times h]}$$

[h×h]

$$\frac{\partial C}{\partial b_2} = \frac{\partial C}{\partial o}\frac{\partial o}{\partial H_2}\frac{\partial H_2}{\partial Z_2}\frac{\partial Z_2}{\partial b_2} = 2(o - o_{des})\overbrace{W_3}^{[h\times 1]} .* \overbrace{\frac{\partial H_2}{\partial Z_2}}^{[h\times 1]}$$

[h×1]

$$\frac{\partial o}{\partial H_2} = W_3$$

$$\frac{\partial H_1}{\partial Z_1} = \begin{cases} 0 & \text{if } H_1(Z_1) = 0 \\ 1 & \text{if } H_1(Z_1) \neq 0 \end{cases}$$

$$\frac{\partial Z_1}{\partial W_1} = i, \quad \frac{\partial Z_1}{\partial b_1} = 1$$

$$\frac{\partial Z_2}{\partial H_1} = W_2$$

$$\frac{\partial C}{\partial W_1} = \frac{\partial C}{\partial o}\frac{\partial o}{\partial H_2}\frac{\partial H_2}{\partial Z_2}\frac{\partial Z_2}{\partial H_1}\frac{\partial H_1}{\partial Z_1}\frac{\partial Z_1}{\partial W_1} = 2(o - o_{des})\left(\left(W_3 .* \frac{\partial H_2}{\partial Z_2}\right)' W_2\right)' .* \frac{\partial H_1}{\partial Z_1} i$$

[h×1]

$$\frac{\partial C}{\partial b_1} = \frac{\partial C}{\partial o}\frac{\partial o}{\partial H_2}\frac{\partial H_2}{\partial Z_2}\frac{\partial Z_2}{\partial H_1}\frac{\partial H_1}{\partial Z_1}\frac{\partial Z_1}{\partial b_1} = 2(o - o_{des})\left(\left(W_3 .* \frac{\partial H_2}{\partial Z_2}\right)' W_2\right)' .* \frac{\partial H_1}{\partial Z_1}$$

[h×1]