Task 4

8 January, 2025 20:09

- $w_{1,3} = 1.5\,;\, w_{2,3} = -2.5\,;\, b_3 = 0.3$
- $w_{1,4} = 1\,;\, w_{2,4} = -2.5\,;\, b_4 = 0.2$
- $w_{3,5} = 4\,;\, w_{4,5} = 3\,;\, b_5 = -0.8$

Forward pass for example 1: $x=(1,2)$; $y=-1$

- Calculate the input to the hidden layer neurons

For n_3: $z_3 = w_{1,3} \cdot x_1 + w_{2,3} \cdot x_2 + b_3 = 1{,}5 \cdot 1 + (-2{,}5) \cdot 2 + 0{,}3 = -3{,}2$

For n_4: $z_4 = w_{1,4} \cdot x_1 + w_{2,4} \cdot x_2 + b_4 = 1 \cdot 1 + (-2{,}5) \cdot 2 + 0{,}2 = 1 - 5 + 0{,}2 = -3{,}8$

- Output from hidden layers

For n_3: $O_3 = \frac{1}{1+e^{3{,}2}} \approx 0{,}0391$

For n_4: $O_4 = \frac{1}{1+e^{3{,}8}} \approx 0{,}0218$

- Calculate input to the output neuron

$z_5 = w_{3,5} \cdot O_3 + w_{4,5} \cdot O_4 + b_5 = 4 \cdot 0{,}0391 + 3 \cdot 0{,}0218 - 0{,}8 = -0{,}5806$

- Calculate output

$O_5 = \frac{1}{1+e^{0{,}5806}} \approx 0{,}3587$

- Calculate error (loss) for example 1

$E_1 = \frac{1}{2}(O_5 - y_1)^2 = \frac{1}{2}(0{,}3587-(-1))^2 = \frac{1}{2} \cdot 1{,}3587^2 \approx 0{,}9230$

Forward pass for example 2: $x=(2,0)$, $y=6$

- Calculate the input to the hidden layer neurons

For n_3: $z_3 = w_{1,3} \cdot x_1 + w_{2,3} \cdot x_2 + b_3 = 1{,}5 \cdot 2 + (-2{,}5) \cdot 0 + 0{,}3 = 3 + 0{,}3 = 3{,}3$

For n_4: $z_4 = w_{1,4} \cdot x_1 + w_{2,4} \cdot x_2 + b_4 = 1 \cdot 2 + (-2{,}5) \cdot 0 + 0{,}2 = 2{,}2$

- Output from hidden layers

For n_3: $O_3 = \frac{1}{1+e^{-3{,}3}} \approx 0{,}9644$

• Output from hidden layers

For n_3: $a_3 = \frac{1}{1+e^{-3,3}} \approx 0,9644$

For n_4: $a_4 = \frac{1}{1+e^{-2,2}} \approx 0,9002$

• Calculate input to the output neuron

$z_5 = w_{3,5} \cdot a_3 + w_{4,5} \cdot a_4 + b_5 = 4 \cdot 0,9644 + 3 \cdot 0,9002 - 0,8 = 5,7582$

• Calculate output

$a_5 = \frac{1}{1+e^{-5,7582}} \approx 0,9968$

• Calculate error (loss) for example 2

$E_2 = \frac{1}{2}(a_5 - y_2)^2 = \frac{1}{2}(0,9968 - 6)^2 \approx 12,5160$

Calculating average error

$E_{avg} = \frac{E_1 + E_2}{2} = 6,7195$

Backpropagation calculations

• Calculate error derivative for the output layer

$\frac{\partial E_1}{\partial a_5} = a_5^{(1)} - y_1 = 0,3587 + 1 = 1,3587$

$\frac{\partial E_2}{\partial a_5} = a_5^{(2)} - y_2 = 0,9968 - 6 = -5,0032$

$\frac{\partial E_{avg}}{\partial a_5} = \frac{1}{2}\left(\left(a_5^{(1)} - y_1\right) + \left(a_5^{(2)} - y_2\right)\right) = \frac{1}{2}\left(\frac{\partial E_1}{\partial a_5} + \frac{\partial E_2}{\partial a_5}\right) = -1,8222$

• Calculate for the hidden layer

$\frac{\partial E_{avg}}{\partial a_3} = \frac{\partial E_{avg}}{\partial a_5} \cdot w_{3,5} = -1,8222 \cdot 4 = -7,2888$

$\frac{\partial E_{avg}}{\partial a_4} = \frac{\partial E_{avg}}{\partial a_5} \cdot w_{4,5} = -1,8222 \cdot 3 = -5,4666$

• Backpropagate to inputs of these neurons (start by calculating sigmoid derivative for each version) $\sigma'(x) = \sigma(x) \cdot (1 - \sigma(x))$

- Backpropagate to inputs of these neurons $\sigma'(x) = \sigma(x)\cdot(1-\sigma(x))$ (derivative for each z version)

$z_3^1 = -3{,}2$ $z_3^2 = 3{,}3$

$z_4^1 = -3{,}8$ $z_4^2 = 2{,}6$

$\sigma'(z_3^1) \approx 0{,}0376$ $\sigma'(z_3^2) \approx 0{,}0343$

$\sigma'(z_4^1) \approx 0{,}0214$ $\sigma'(z_4^2) \approx 0{,}0898$

$\frac{\partial E_{avg}}{\partial z_3^1} = -7{,}2888 \cdot 0{,}0376 = -0{,}2966$ $\frac{\partial E_{avg}}{\partial z_4^1} = -5{,}4666 \cdot 0{,}0214 = -0{,}1169$

$\frac{\partial E_{avg}}{\partial z_3^2} = -7{,}2888 \cdot 0{,}0343 = -0{,}2705$ $\frac{\partial E_{avg}}{\partial z_4^2} = -5{,}4666 \cdot 0{,}0898 = -0{,}4909$

$\frac{\partial E_{avg}}{\partial z_3} \approx -0{,}2835$ $\frac{\partial E_{avg}}{\partial z_4} \approx -0{,}3039$

- Gradients for weights and biases

$\frac{\partial E_{avg}}{\partial E_{w_{1,3}}} = -0{,}2835 \cdot a_1 = -0{,}4252$ $\frac{\partial E_{avg}}{\partial w_{2,3}} = -0{,}2835 \cdot a_2 = -0{,}2835$

$\frac{\partial E_{avg}}{\partial E_{w_{1,4}}} = -0{,}3039 \cdot a_1 = -0{,}3058$ $\frac{\partial E_{avg}}{\partial w_{2,4}} = -0{,}3039 \cdot a_2 = -0{,}3039$

$\frac{\partial E_{avg}}{\partial b_3} = -0{,}2835$ $\frac{\partial E_{avg}}{\partial b_4} = -0{,}3039$

Calculating new weights

$w_{i,j} = w_{i,j} - \eta \cdot \frac{\partial E_{avg}}{\partial w_{i,j}}$

$b_i = b_i - \eta \cdot \frac{\partial E_{avg}}{\partial b_i}$

where η is the learning rate. We will be using 0,1.

Old values:

- $w_{1,3} = 1.5$; $w_{2,3} = -2.5$; $b_3 = 0.3$
- $w_{1,4} = 1$; $w_{2,4} = -2.5$; $b_4 = 0.2$

- $w_{1,4} = 1\,; w_{2,4} = -2.5\,; b_4 = 0.2$
- $w_{3,5} = 4\,; w_{4,5} = 3\,; b_5 = -0.8$

$w_{1,3} = 1{,}54252$ $w_{2,3} = -2{,}47165$ $b_3 = 0{,}32835$

$w_{1,4} = 1{,}03058$ $w_{2,4} = -2{,}46941$ $b_4 = 0{,}23039$

$w_{3,5} = 4{,}0358$ $w_{4,5} = 3{,}03039$ $b_5 = -0{,}7761$

We now have new weights, which we can use to do a forward step

Another Forward pass to check changes in predictions

$z_3 \approx -3{,}0124$

$a_3 \approx 0{,}0468$

$z_4 \approx -3{,}6778$

$a_4 \approx 0{,}0246$

$z_5 \approx -0{,}51227$

$a_5 \approx -0{,}51227$

$E_1 \approx 0{,}11894$

$z_3 \approx 3{,}4735$

$a_3 \approx 0{,}9701$

$z_4 \approx 2{,}2815$

$a_4 \approx 0{,}9085$

$z_5 \approx 6{,}5106$

$a_5 \approx 6{,}5106$

$E_2 \approx 0{,}13035$

$E_{avg} = 0{,}124645$