*Colin Kirby*
*co201115@ucf.edu*

**Question # 1 : Compute the output of the following neural network, when :**

*Inputs* : $x_1 = 1$ , $x_2 = 2$

*Weights (Input to Hidden)* : $w_1 = 0.2$ , $w_2 = 0.1$ , $w_3 = 0.3$ , $w_4 = 0.2$

*Weights (Hidden to Output)* : $w_5 = 0.1$ , $w_6 = 0.2$ , $w_7 = 0.4$ , $w_8 = 0.3$

*Biases* : $b_1 = 0.1$ , $b_2 = 0.2$

**The activation functions of the neurons are both ReLU functions. Provide all steps. You may use python to do the step-by-step computation.**

For the first two neurons, we'll define them as $h_1$ and $h_2$ with inputs $x_1 = 1$ and $x_2 = 2$ .

*Neuron #1 ( $h_1$ ) :*

$h_1$ has inputs $b_1 = 0.1$ , $w_1 = 0.2$ , $w_2 = 0.1$ .

$$h_1 = ReLU\ (\ w_1 x_1 + w_2 x_2 + b_1\ )$$

$$h_1 = ReLU\ (\ (\ 0.2\ *\ 1\ ) + (\ 0.1\ *\ 2\ ) + 0.1\ )$$

$$h_1 = ReLU\ (\ 0.2 + 0.2 + 0.1\ ) = ReLU\ (\ 0.5\ )$$

$$h_1 = 0.5$$

*Neuron #2 ( $h_2$ ) :*

$h_2$ has inputs $b_1 = 0.1$ , $w_3 = 0.3$ , $w_4 = 0.2$ .

$$h_2 = ReLU\ (\ w_3 x_1 + w_4 x_2 + b_2\ )$$

$$h_2 = ReLU\ (\ \ (\ 0.3 * 1\ )\ ) + (0.2 * 2\ ) + 0.1\ )$$

$$h_2 = ReLU\ (\ 0.3 + 0.4 + 0.2\ ) = ReLU\ (\ 0.8\ )$$

$$h_2 = 0.8$$

Moving onto the output neurons, we now have the inputs $h_1 = 0.5$ and $h_2 = 0.8$ .

*Output Neuron #1 ( $o_1$ ) :*

$o_1$ has inputs $b_2 = 0.2$ , $w_5 = 0{,}1$ , $w_6 = 0.2$ .

$$o_1 = ReLU\ (\ w_5 h_1 + w_6 h_2 + b_2)$$

$$o_1 = ReLU\ (\ (\ 0.1 * 0.5\ ) + (\ 0.2 * 0.8\ ) + 0.2\ )$$

$$o_1 = ReLU\ (\ 0.05 + 0.16 + 0.2\ ) = ReLU\ (\ 0.41\ )$$

$$o_1 = 0.41$$

*Colin Kirby*
*co201115@ucf.edu*

*Output Neuron #2 ( $o_2$ ) :*

$o_2$ has inputs $b_2 = 0.2$ , $w_7 = 0.4$ , $w_8 = 0.3$ .

$$o_2 = ReLU\ (\ w_7 h_1 + w_8 h_2 + b_2)$$

$$o_2 = ReLU\ (\ (\ 0.4 * 0.5\ ) + (\ 0.3 * 0.8\ ) + 0.2\ )$$

$$o_2 = ReLU\ (\ 0.2 + 0.24 + 0.2\ ) = ReLU\ (\ 0.64\ )$$

$$o_2 = 0.64$$

*Final Outputs ( $o_1$ and $o_2$ ) :*

$$o_1 = 0.41$$

$$o_2 = 0.64$$

*Colin Kirby*
*co201115@ucf.edu*

**Question # 2 : Assume the ground truth in the training for input $(\boldsymbol{x_1}, \boldsymbol{x_2}) = (\mathbf{1}, \mathbf{2})$ is $(\boldsymbol{o_1}, \boldsymbol{o_2}) = (\mathbf{0}, \mathbf{1})$ , the current w's are the values in Question 1, and the loss function is defined as :**

$$L = (o_1 - \widehat{o_1})^2 + (o_2 - \widehat{o_2})^2$$

Compute $\frac{\partial L}{\partial w_3}$ and $\frac{\partial L}{\partial w_7}$ . Provide all steps. You may use Python to do step-by-step computation.

We have the values $h_1 = 0.5$ , $h_2 = 0.8$ and $o_1 = 0.41$ , $o_2 = 0.64$ from the previous question.

With the Ground Truth Values $o_1 = 0$ and $o_2 = 1$ . We can now plug into the Loss function :

$$L = (0.41 - 0)^2 + (0.64 - 1)^2$$

$$L = (0.41)^2 + (-0.36)^2$$

$$L = 0.1681 + 0.1296$$

$$L = 0.2977$$

We need to solve these gradients :

$$\frac{\partial L}{\partial w_7} = \frac{\partial L}{\partial o_2} * \frac{\partial o_2}{\partial w_7}$$

$$\frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial h_2} * \frac{\partial h_2}{\partial w_3}$$

Now to compute the necessary values for the gradients.

$$\frac{\partial L}{\partial o_1} = 2\,(\,o_1 - \widehat{o_1}) = 2\,(\,0.41 - 0\,) = 0.82$$

$$\frac{\partial L}{\partial o_2} = 2\,(\,o_2 - \widehat{o_2}) = 2\,(\,0.64 - 1\,) = -0.72$$

Gradient of Output with respect to Weight for $w_7$ :

$$\frac{\partial o_2}{\partial w_7} = \; w_7 h_1 + w_8 h_2 + b_2 \rightarrow h_1 = 0.5$$

From these values we can solve for $\frac{\partial L}{\partial w_7}$ :

$$\frac{\partial L}{\partial w_7} = \frac{\partial L}{\partial h_2} * \frac{\partial h_2}{\partial w_3} = -0.72 * 0.5 = \; -0.36$$

Now to solve for the remaining values.

*Colin Kirby*
*co201115@ucf.edu*

Solving for $\frac{\partial L}{\partial h_2}$:

$$\frac{\partial L}{\partial h_2} = \left( w_6 * \frac{\partial L}{\partial o_1} \right) + \left( w_8 * \frac{\partial L}{\partial o_2} \right) = (\ 0.2 * 0.82\ ) + (0.3 * -0.72\ )$$

$$\frac{\partial L}{\partial h_2} = 0.164 - 0.216 = -\ 0.052$$

And finally solving for $\frac{\partial h_2}{\partial w_3}$:

$$\frac{\partial h_2}{\partial w_3} = \ w_3 x_1 + w_4 x_2 + b_2 \rightarrow x_1 = 1$$

Now we can solve for $\frac{\partial L}{\partial w_3}$:

$$\frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial h_2} * \frac{\partial h_2}{\partial w_3} = \ -0.052 * 1 = \ -0.052$$

Final Answers ( $\frac{\partial L}{\partial w_3}$ and $\frac{\partial L}{\partial w_7}$ ) :

$$\frac{\partial L}{\partial w_7} = \ -0.36$$

$$\frac{\partial L}{\partial w_3} = \ -0.052$$

*Colin Kirby*
*co201115@ucf.edu*

**Question #3 : Now we add a softmax output lay for multi-class classification. Assume there are two classes – class 1 and class 2. The ground truth for input $(x_1, x_2) = (1, 2)$ is class 2, the current w's are the values in Question #1 and the loss function is cross-entropy. Compute $\frac{\partial L}{\partial w_3}$ and $\frac{\partial L}{\partial w_7}$. Provide all steps. You may use Python to do the step-by-step computation.**

Considering our calculated and given values here :

$$h_1 = 0.5\,, h_2 = 0.8$$

$$o_1 = 0.41\,, o_2 = 0.64$$

Lets now apply the softmax function of the exp over the sum of exps.

$$e^{z_1} = e^{0.41} \cong 1.506$$

$$e^{z_2} = e^{0.64} \cong 1.896$$

$$\sum e^{z} = 1.506 + 1.896 = 3.402$$

The softmax outputs are then :

$$\widehat{y_1} = \frac{1.506}{3.402} \cong 0.443$$

$$\widehat{y_2} = \frac{1.896}{3.402} \cong 0.557$$

In the question it is stated that the ground truth is class 2 so we can tell $y_{true} = [0, 1]$.

Utilizing this for our gradient of loss with respect to our previously calculated softmax outputs is :

$$\frac{\partial L}{\partial z_i} = \widehat{y_i} - y_{true,i}$$

For $z_1$ , $\frac{\partial L}{\partial z_1} = 0.443 - 0 = 0.443$

For $z_2$ , $\frac{\partial L}{\partial z_2} = 0.557 - 1 = -0.443$

We now can solve for the $\frac{\partial L}{\partial w_7}$ .

$$\frac{\partial L}{\partial w_7} = \frac{\partial L}{\partial z_2} * \frac{\partial z_2}{\partial w_7} = (\,-0.443\,) * (\,0.5\,) = -0.222$$

*Colin Kirby*
*co201115@ucf.edu*

Now we need to compute $\frac{\partial L}{\partial h_2}$ to properly solve for $\frac{\partial L}{\partial w_3}$.

$$\frac{\partial L}{\partial h_2} = \mathrm{w}_6 \frac{\partial L}{\partial z_1} + w_8 \frac{\partial L}{\partial z_2} = (0.2 * 0.443) + (0.3 * \ -0.443)$$

$$\frac{\partial L}{\partial h_2} = 0.00886 - 0.1329 = -0.0443$$

Now we can solve for $\frac{\partial L}{\partial w_3}$ using our previous value in $\frac{\partial h_2}{\partial w_3}$ from Question #2.

$$\frac{\partial L}{\partial w_3} = \ \frac{\partial L}{\partial h_2} * \frac{\partial h_2}{\partial w_3} = \ -0.0443 * x_1 = \ -0.0443 * 1 = \ -0.0443$$

Final Answers ( $\frac{\partial L}{\partial w_3}$ and $\frac{\partial L}{\partial w_7}$ ) :

$$\frac{\partial L}{\partial w_7} = \ -0.222$$

$$\frac{\partial L}{\partial w_3} = \ -0.0443$$