Now we calculate the derivative of output with respect to input for the output layer

$$\frac{\partial O_{out1}}{\partial O_{in1}} = \frac{\partial\left(e^{O_{in1}}/(e^{O_{in1}}+e^{O_{in2}}+e^{O_{in3}})\right)}{\partial O_{in1}}$$

We know that $\frac{\partial O_{out1}}{\partial O_{in1}} = \left(e^{O_{in1}}(e^{O_{in2}}+e^{O_{in3}})\right)/(e^{O_{in1}}+e^{O_{in2}}+e^{O_{in3}})^2$

$$\therefore \begin{bmatrix} \frac{\partial O_{out1}}{\partial O_{in1}} \\ \frac{\partial O_{out2}}{\partial O_{in2}} \\ \frac{\partial O_{out3}}{\partial O_{in3}} \end{bmatrix} = \begin{bmatrix} \left(e^{O_{in1}}(e^{O_{in2}}+e^{O_{in3}})\right)/(e^{O_{in1}}+e^{O_{in2}}+e^{O_{in3}})^2 \\ \left(e^{O_{in2}}(e^{O_{in1}}+e^{O_{in3}})\right)/(e^{O_{in1}}+e^{O_{in2}}+e^{O_{in3}})^2 \\ \left(e^{O_{in3}}(e^{O_{in1}}+e^{O_{in2}})\right)/(e^{O_{in1}}+e^{O_{in2}}+e^{O_{in3}})^2 \end{bmatrix} = \begin{bmatrix} e^{1.8658}\left(e^{2.2292}+e^{2.8204}\right)/(e^{1.8658}+e^{2.2292}+e^{2.8204})^2 \\ e^{2.2292}\left(e^{1.8658}+e^{2.8204}\right)/(e^{1.8658}+e^{2.2292}+e^{2.8204})^2 \\ e^{2.8204}\left(e^{1.8658}+e^{2.2292}\right)/(e^{1.8658}+e^{2.2292}+e^{2.8204})^2 \end{bmatrix}$$

($e^{1.8658} \to 6.4611$, $e^{2.2292} \to 9.2924$, $e^{2.8204} \to 16.7836$)

$$\begin{bmatrix} \frac{\partial O_{out1}}{\partial O_{in1}} \\ \frac{\partial O_{out2}}{\partial O_{in2}} \\ \frac{\partial O_{out3}}{\partial O_{in3}} \end{bmatrix} = \begin{bmatrix} (6.4611)(26.0759)/(32.5371)^2 \\ (9.2924)(25.2447)/(32.5371)^2 \\ (16.7836)(15.7535)/(32.5371)^2 \end{bmatrix} = \begin{bmatrix} (168.4789)/(32.5371)^2 \\ (214.4091)/(32.5371)^2 \\ (264.4004)/(32.5371)^2 \end{bmatrix} = \begin{bmatrix} 0.1591 \\ 0.2025 \\ 0.2497 \end{bmatrix}$$

($(32.5371)^2 \to 1{,}058.6629$)

Now we proceed to calculate derivative for each input neuron with respect to each weight.

$$\frac{\partial O_{in1}}{\partial W_{K1L1}} = \frac{\partial\left[(h_{2out1} * W_{K1L1}) + (h_{2out2} * W_{K2L1}) + (h_{3out3} * W_{K3L1})\right]}{\partial W_{K1L1}}$$

$$\frac{\partial O_{in1}}{\partial W_{K1L1}} = \frac{\partial(h_{2out1} * W_{K1L1})}{\partial W_{K1L1}} = h_{2out1}$$

Likewise in similar fashion, we can get the others as:

$$\begin{bmatrix} \frac{\partial O_{in1}}{\partial W_{K1L2}} \\ \frac{\partial O_{in2}}{\partial W_{K1L2}} \\ \frac{\partial O_{in3}}{\partial W_{K1L2}} \end{bmatrix} = \begin{bmatrix} h_{2out1} \\ h_{2out2} \end{bmatrix} \begin{bmatrix} \frac{\partial O_{in3}}{\partial W_{K1L3}} \\ \frac{\partial O_{in3}}{\partial W_{K1L3}} \end{bmatrix} = \begin{bmatrix} h_{2out1} \\ h_{2out2} \end{bmatrix} \begin{bmatrix} h_{2out1} \\ h_{2out2} \\ h_{2out3} \end{bmatrix} = \begin{bmatrix} 0.938 \\ 0.94 \\ 0.98 \end{bmatrix}$$