Manual Calculation of DL

In part one of the deep learning calculation, we start by calculating the gradients for the output layer weights using this equation:

❶ Every "z" in this formula is the actual output for **node c1.**

$$\frac{\partial E}{\partial W_5} = (z-t)\, z(1-z)\, out_{b1}$$

❷ This "t" is the target output for **node c1.**

Using the equation, I was able to calculate weights 5, 6, 7, and 8.

<u>Part 2:</u> In part two of the deep learning calculation process, we would calculate the gradients for the output layer bias weights

$$\delta_z = (z-t)\, z(1 - z)$$

We were able to calculate bias weights 3, 4, and 2

<u>Part 3:</u> calculating gradients for hidden layer weights

$$\frac{\partial E}{\partial W_1} = \Big(\overbrace{\sum_c \delta_z W_i}^{❶}\Big)\underbrace{out_{b1}(1 - out_{b1})}_{❷}\overset{❸}{out_{a1}}$$

This equation is a summary of the previous equations that we have calculated. Doing this portion of the calculation process is easy.

<u>Part 4</u>: calculating gradients for hidden layer bias weights.

$$\delta_b = \Big(\sum_c \delta_z W_i\Big) out_i (1 - out_i)$$

❶ All of the letter "i"'s refer to a unique value. This value depends on the gradient we are calculating.

<u>Part 5</u>: The final step in backpropagation is updating the weights

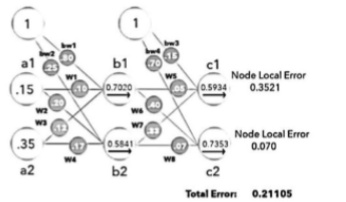
❶ The old weight **w8**, which is being updated.

❸ The partial derivative of **the total error** with respect to the weight **w8.**

$$W_{8_{new}} = W_8 - \eta * \frac{\partial E}{\partial W_8}$$

❷ The greek letter eta represents the learning rate.

Lab 1 Calculating Gradients



$$\frac{\partial E}{\partial W_5} = (z-t)\, z(1-z)\, out_{b1}$$

1 Every "z" in this formula is the actual output for **node c1**.

2 This "t" is the target output for **node c1**.

$\frac{\partial E}{\partial W_5} = (0.5934 - 0) * 0.5934(1 - 0.5934) * 0.7020 = \underline{0.1005}$

$\frac{\partial E}{W_6} = (0.7353 - 1) * 0.7353(1 - 0.7353) * 0.7020 = \underline{-0.0362}$

$\frac{\partial E}{W_7} = (0.5934 - 0)\, 0.5934(1 - 0.5934) * 0.5841 = \underline{0.0836}$

$\frac{\partial E}{W_8} = (0.7353 - 1)\, 0.7353(1 - 0.7353) * 0.5841 = \underline{-0.0301}$

Bias Weights

$$\boxed{\delta_z = (z-t)\, z(1-z)}$$

Delta z

$BW_3 = (0.5934 - 0) * .5934(1 - .5934) = \underline{0.1432}$

$BW_4 = (0.7353 - 1) * .7353(1 - .7353) = \underline{-0.0515}$

Hidden Layer

$$\boxed{\frac{\partial E}{\partial W_1} = \left(\sum_c \delta_z W_i\right) out_{b1}(1 - out_{b1})\, out_{a1}}$$

① $W_1 \rightarrow$ $0.1432 * \overset{W_5}{0.05} = 0.0072$

$-0.0515 * \overset{W_6}{0.4} = -0.0206$

$0.0072 - 0.0206 = -0.0134$

$\sum \delta_z W_i = \underline{-0.0134}$

$out_{b_1}(1 - out_{b_1}) \rightarrow 0.7020 * (1 - 0.7020) = \underline{0.2092}$

$-0.0134 * .2092 * 0.15 = \underline{-0.00042}$

$W_2 \rightarrow$ $\delta_z = \overset{c_1}{0.1432} * \overset{W_7}{.33} = \underline{0.0473}$

$\delta_z = -0.0515 * \overset{W_8}{.07} = \underline{-0.0036}$

$\underline{0.0437}$

② $out_{b_2}(1 - out_{b_2})$

$0.5841 * (1 - 0.5841) = \underline{0.2429}$

③ $0.0437 * 0.2429 * .15 = 0.00159$

$W_3 \rightarrow -0.0134 * 0.2092 * 0.35 = -0.00098$

$W_4 \rightarrow 0.0437 * 0.2429 * 0.35 = 0.00372$

Hidden Layer Bias Weights

$$\delta_b = \left(\sum_c \delta_z W_i\right) out_i (1 - out_i)$$

❶ All of the letter "i"'s refer to a unique value. This value depends on the gradient we are calculating.

$$\frac{\partial E}{\partial BW_1} = \left(\Sigma \delta W_{5+6}\right) out_{b_1} (1 - out_{b_1})$$

$$\frac{\partial E}{BW_2} = \left(\Sigma \delta W_{7+8}\right) out_{b_2} (1 - out_{b_2})$$

$(-0.0134) * 0.7020\ (1 - 0.7020) = \underline{-0.0028}$

$0.0437 * 0.5841\ (1 - 0.5841) = \underline{0.0106}$

Updating Weights

❶ The old weight **w8**, which is being updated.

❸ The partial derivative of **the total error** with respect to the weight **w8.**

$$W8_{new} = W8 - \eta * \frac{\partial E}{\partial W8}$$

❷ The greek letter eta represents the learning rate.

$B_{W_4} = 0.7 - 0.5 * -0.0515 = 0.7258$

$B_{W_3} = 0.15 - 0.5 * 0.1432 = 0.0784$

$B_{W_2} = 0.25 - 0.5 * 0.016 = 0.242$

$B_{W_1} = 0.8 - 0.5 * -0.0028 = 0.814$

$W_8 = 0.07 - 0.5 * (-0.0301) = 0.0851$

$W_7 = 0.33 - 0.5 * .0836 = 0.2882$

$W_6 = .4 - 0.5 * -0.0362 = 0.4181$

$W_5 = 0.05 - 0.5 * 0.1005 = -0.00025$

$W_4 = 0.17 - 0.5 * 0.00372 = 0.1514$

$W_3 = 0.12 - 0.5 * -0.00098 = 0.1205$

$W_2 = 0.2 - 0.5 * 0.00159 = 0.1992$

$W_1 = 0.1 - 0.5 * -0.000042 = 0.1002$