

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:

$$\frac{\partial E}{\partial W_5} = (z-t) z(1-z) \text{out}_{b1}$$

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

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

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

Part 2: 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

Part 3: calculating gradients for hidden layer weights

$$\frac{\partial E}{\partial W_1} = \left(\sum_c \delta_z W_i \right) \underbrace{\text{out}_{b1} (1 - \text{out}_{b1})}_{\text{②}} \underbrace{\text{out}_{a1}}_{\text{③}}$$

①

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

Part 4: calculating gradients for hidden layer bias weights.

$$\delta_b = \left(\sum_c \delta_z W_i \right) \text{out}_i (1 - \text{out}_i)$$

① All of the letter "i"s refer to a unique value. This value depends on the gradient we are calculating.

Part 5: 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**.

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

❷ The greek letter eta represents the learning rate.

Hidden Layer Bias Weights

$$\delta_i = (\sum_c \delta_c W_{ci}) \text{out}_i (1 - \text{out}_i)$$

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

$$\frac{\partial E}{\partial BW_1} = (\sum_c \delta_c W_{c1}) \text{out}_{b_1} (1 - \text{out}_{b_1})$$

$$(-0.0134) * 0.7020 (1 - 0.7020) = -0.0028$$

$$\frac{\partial E}{\partial BW_2} = (\sum_c \delta_c W_{c2}) \text{out}_{b_2} (1 - \text{out}_{b_2})$$

$$0.0437 * 0.5841 (1 - 0.5841) = 0.0106$$

Updating Weights

The old weight w_8 , which is being updated.

The partial derivative of the total error with respect to the weight w_8 .

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

The greek letter eta represents the learning rate.

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

$$W_7 = 0.33 - 0.5 * 0.0836 = 0.2882$$

$$W_6 = 0.4 - 0.5 * -0.0362 = 0.4181$$

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

$$W_4 = 0.17 - 0.5 * 0.00372 = 0.1514$$

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

$$W_2 = 0.2 - 0.5 * 0.00159 = 0.1992$$

$$W_1 = 0.1 - 0.5 * -0.00042 = 0.1002$$

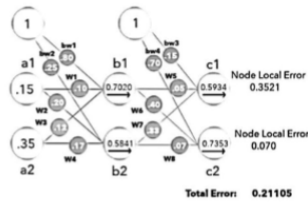
$$BW_4 = 0.7 - 0.5 * -0.0515 = 0.7258$$

$$BW_3 = 0.15 - 0.5 * 0.1432 = 0.0784$$

$$BW_2 = 0.25 - 0.5 * 0.016 = 0.242$$

$$BW_1 = 0.8 - 0.5 * -0.0028 = 0.814$$

Lab 1 Calculating Gradients



$$\frac{\partial E}{\partial W_i} = (z - t) z (1 - z) \text{out}_{b1}$$

Every "z" in this formula is the actual output for node c1.
This "t" is the target output for node c1.

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

$$\frac{\partial E}{\partial W_6}$$

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

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

$$\frac{\partial E}{\partial W_8}$$

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

Bias Weights

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

Delta z

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

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

Hidden Layer

$$\frac{\partial E}{\partial W_i} = \left(\sum_c \delta_c W_i \right) \text{out}_{b1} (1 - \text{out}_{b1}) \text{out}_{a1}$$

$$\begin{aligned} W_1 &\rightarrow 0.1432 * 0.05 = 0.0072 \\ &- 0.0515 * 0.4 = -0.0206 \end{aligned} \quad \left. \begin{aligned} &0.0072 - 0.0206 = -0.0134 \\ &\sum \delta_c W_i = -0.0134 \end{aligned} \right\}$$

$$\text{out}_{b2} (1 - \text{out}_{b2}) \rightarrow 0.7020 * (1 - 0.7020) = \underline{0.2092}$$

$$-0.0134 * 0.2092 * 0.15 = \underline{-0.00042}$$

$$\begin{aligned} W_2 &\rightarrow \delta_2 = 0.1432 * .33 = \underline{0.0473} \\ &\delta_2 = -0.0515 * .07 = \underline{-0.0036} \end{aligned} \quad \left. \begin{aligned} &0.0473 \\ &-0.0036 \end{aligned} \right\} \underline{0.0437}$$

$$\begin{aligned} \text{out}_{b2} (1 - \text{out}_{b2}) \\ &0.5841 * (1 - 0.5841) = \underline{0.2429} \end{aligned}$$

$$\textcircled{3} 0.0437 * 0.2429 * .15 = \underline{0.00159}$$

$$W_3 \rightarrow -0.0134 * 0.2092 * 0.35 = \underline{-0.00098}$$

$$W_4 \rightarrow 0.0437 * 0.2429 * 0.35 = \underline{0.00372}$$