13_Backpropagation_3

30 August 2024 11:09

In general,

$$\delta_i^{l} = \left(\sum_{j=1}^{b_{l+1}} \delta_j^{l+1} \cdot w_{ji}^{l}\right) \cdot f'(z_i^{l})$$

\# of neurons in $(n+1)^{th}$ layer.

For biases,

$$\frac{\partial L}{\partial b_i^3} = (\hat{y}_i - y_i)\,\frac{\partial \hat{y}_i}{\partial b_i^3}$$

$$\hat{y}_i = a_i^4$$

$$\underbrace{\frac{\partial a_i^4}{\partial z_i^4}}_{f'(z_i^4)} \cdot \underbrace{\frac{\partial z_i^4}{\partial b_i^3}}_{1}$$

$$\Rightarrow \quad \frac{\partial L}{\partial b_i^3} = (\hat{y}_i - y_i)\, f'(z_i^4)$$

$$\delta_i^{l} = \sum_{j=1}^{b_{l+1}} \left(\delta_j^{l+1} w_{ji}^{l}\right) f'(z_i^{l})$$

$j=1$

$$\delta_1^3 = \sum_{j=1}^{2} \delta_1^4 \cdot W_{ji}^3 \cdot f'(z_1^3)$$

$$= \left(\delta_1^4 W_{11}^3 + \delta_2^4 W_{21}^3\right) \cdot f'(z_1^3)$$

$$\delta_2^3 = \left(\delta_1^4 W_{12}^3 + \delta_2^4 W_{22}^3\right) \cdot f'(z_2^3)$$

$$\delta_3^3 = \left(\delta_1^4 W_{13}^3 + \delta_2^4 \cdot W_{23}^3\right) \cdot f'(z_3^3)$$

$$\begin{bmatrix} \delta_1^3 \\ \delta_2^3 \\ \delta_3^3 \end{bmatrix} = \begin{bmatrix} W_{11}^3 & W_{21}^3 \\ W_{12}^3 & W_{22}^3 \\ W_{13}^3 & W_{23}^3 \end{bmatrix} \begin{bmatrix} \delta_1^4 \\ \delta_2^4 \end{bmatrix} \odot \begin{bmatrix} f'(z_1^3) \\ f'(z_2^3) \\ f'(z_3^3) \end{bmatrix}$$

element - wise multiplication

$$W^3 = \begin{bmatrix} W_{11}^3 & W_{12}^3 & W_{13}^3 \\ W_{21}^3 & W_{22}^3 & W_{23}^3 \end{bmatrix}$$

$$\delta^3 = \left[(W^3)^T \cdot \delta^4\right] \odot f'(z^3)$$

Delta Learning Rule

$$\boxed{\delta^l = \left((W^l)^T \delta^{l+1}\right) . f'(z^l)}$$