

$$1) z_j^l = \sum_k a_j^{l-1} w_{kj} + b_j^l$$

$$(a) a_j^l = f(z_j^l) \quad \frac{da_j^l}{dz_j^l} = f'(z_j^l)$$

$$\Delta_j^l = \frac{dC}{dz_j^l} = \frac{dC}{da_j^l} \frac{da_j^l}{dz_j^l} = \frac{dC}{da_j^l} f'(z_j^l)$$

for  $l < L$

$$\begin{aligned} \Delta_j^l &= \frac{dC}{dz_j^l} = \sum_k \frac{dC}{dz_k^{l+1}} \frac{dz_k^{l+1}}{da_j^l} \frac{da_j^l}{dz_j^l} \\ &= \sum_k \Delta_k^{l+1} [\omega^{(l+1)^T}]_{kj} f'(z_j^l) \end{aligned}$$

$$(b) \frac{dC}{dw_{ij}^l} = \frac{dC}{dz_j^l} \frac{dz_j^l}{dw_{ij}^l} = \Delta_j^l a_i^{l-1} \frac{dw_{ij}^l}{dw_{ij}^l} = \Delta_j^l a_i^{l-1}$$

$$\frac{dC}{db_i^l} = \frac{dC}{dz_j^l} \frac{dz_j^l}{db_i^l} = \Delta_j^l$$

$$2) a_j^l = z_j^l$$

$$\Delta_j^l = \frac{dC}{dz_j^l} = \frac{dC}{da_j^l} \frac{da_j^l}{dz_j^l} = \frac{dC}{da_j^l}$$

$l < L$

$$\Delta_j^l = \frac{dC}{dz_j^l} = \sum_k \frac{dC}{dz_k^{l+1}} \frac{dz_k^{l+1}}{da_j^l} \frac{da_j^l}{dz_j^l} = \sum_k \Delta_k^{l+1} [\omega^{(l+1)^T}]_{kj}$$

$$\frac{dC}{dw_{ij}^l} = a_j^{l-1} \Delta_j^l \quad \frac{dC}{db_i^l} = \Delta_j^l$$