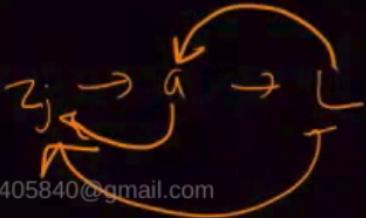


$$\frac{\partial L}{\partial b_j} = \delta_j^L \cdot 1$$

rohit58405840@gmail.com



For Output layer

where -

$$① \frac{\partial L}{\partial w_{ij}} = \delta_j^L \cdot a_i^{L-1}$$

$$\delta_j^L = (a_j - y_j) \odot \sigma'(z_j)$$

$$② \frac{\partial L}{\partial b_j} = \delta_j^L \cdot 1$$





$$\frac{\partial L}{\partial b_j} = \delta_{j+1} \cdot 1$$

For output layer

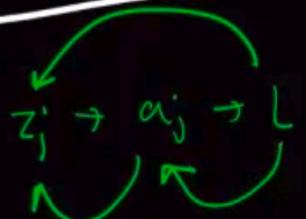
$$\frac{\partial L}{\partial z_j} = \frac{\partial L}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_j}$$

where -

$$\delta_j^L = (a_j - y_j) \odot \sigma'(z_j)$$

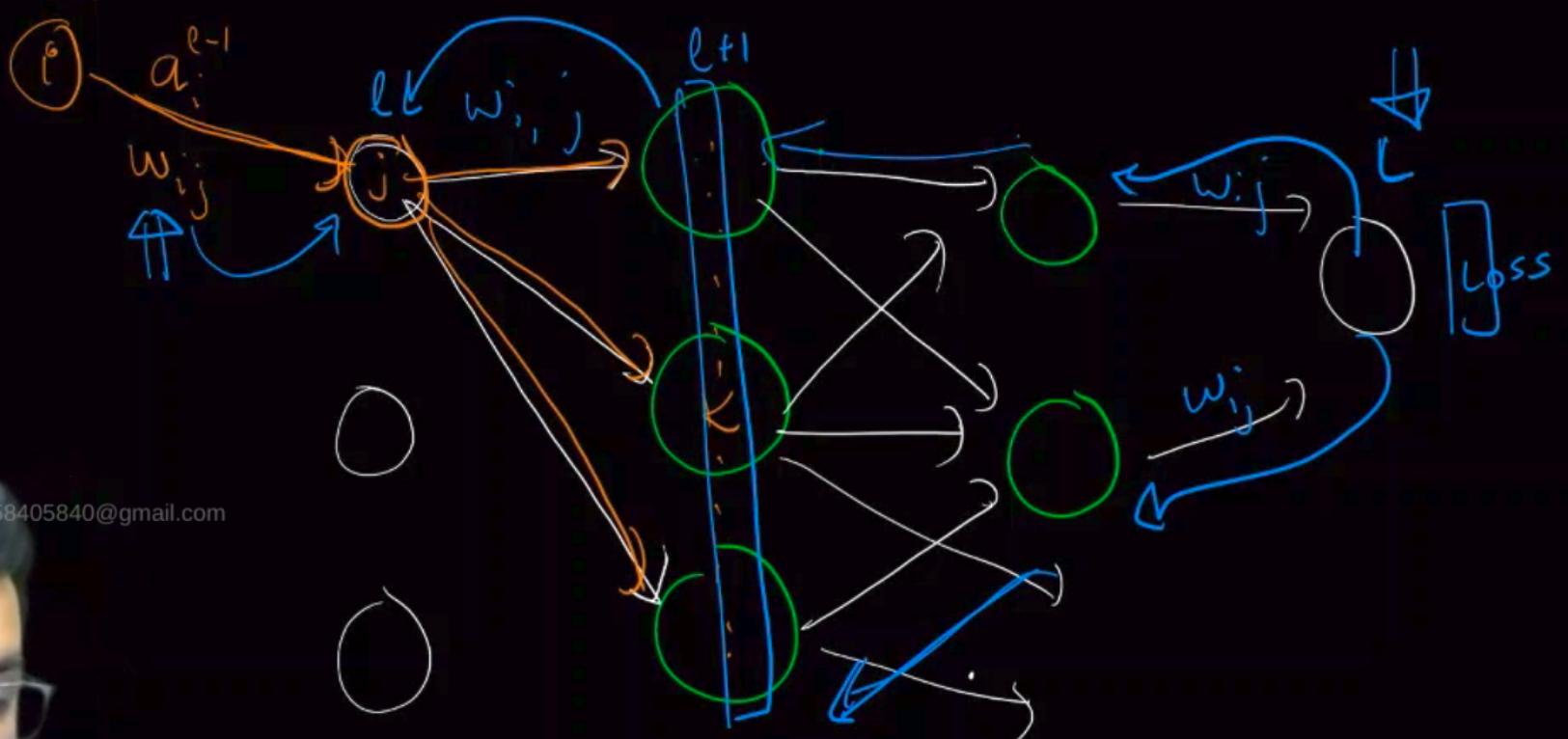
①  $\frac{\partial L}{\partial w_{ij}} = \delta_j^L \cdot a_i^{L-1}$

②  $\frac{\partial L}{\partial b_j} = \delta_j^L \cdot 1$



rohit58405840@gmail.com





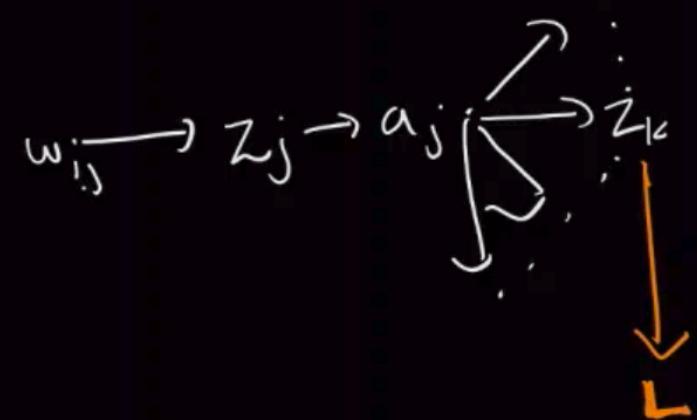
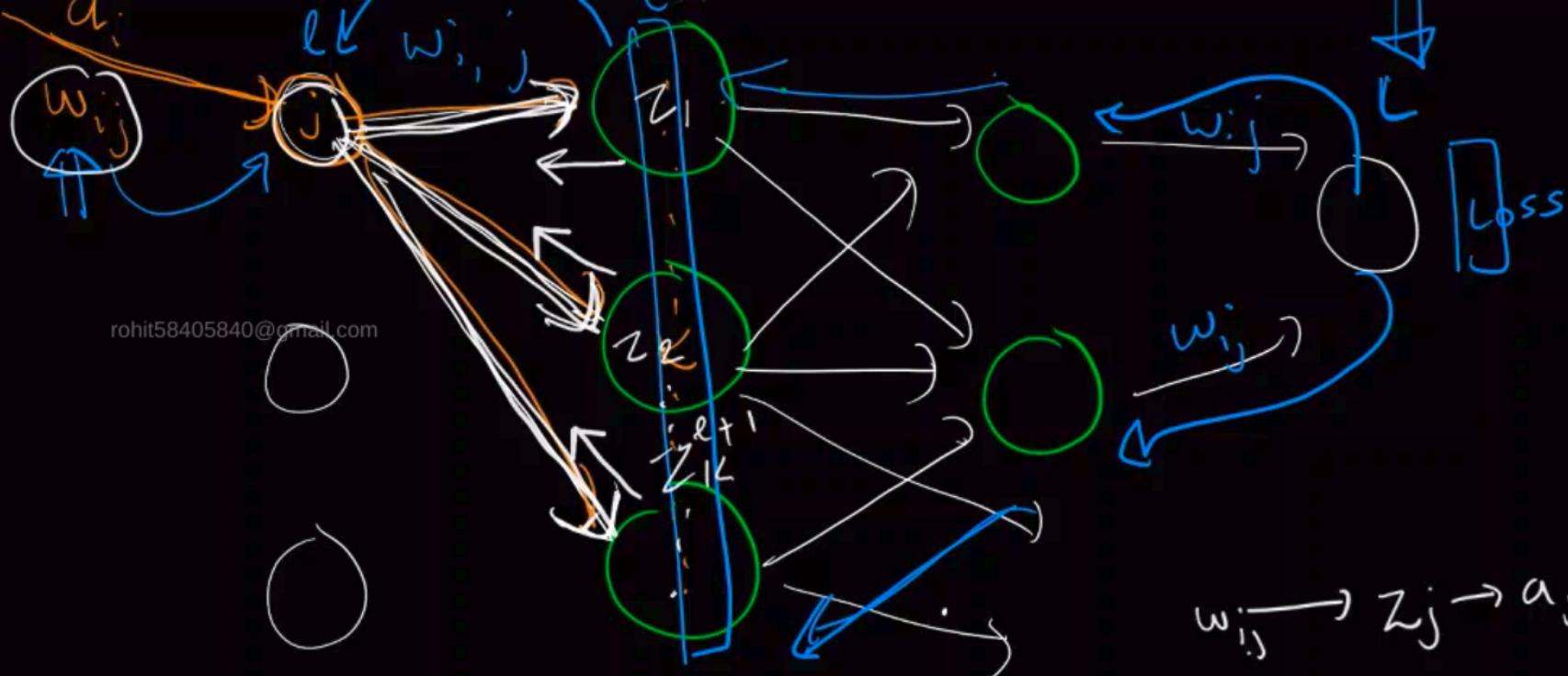
rohit58405840@gmail.com

$$\frac{\partial L}{\partial w_{ij}}$$

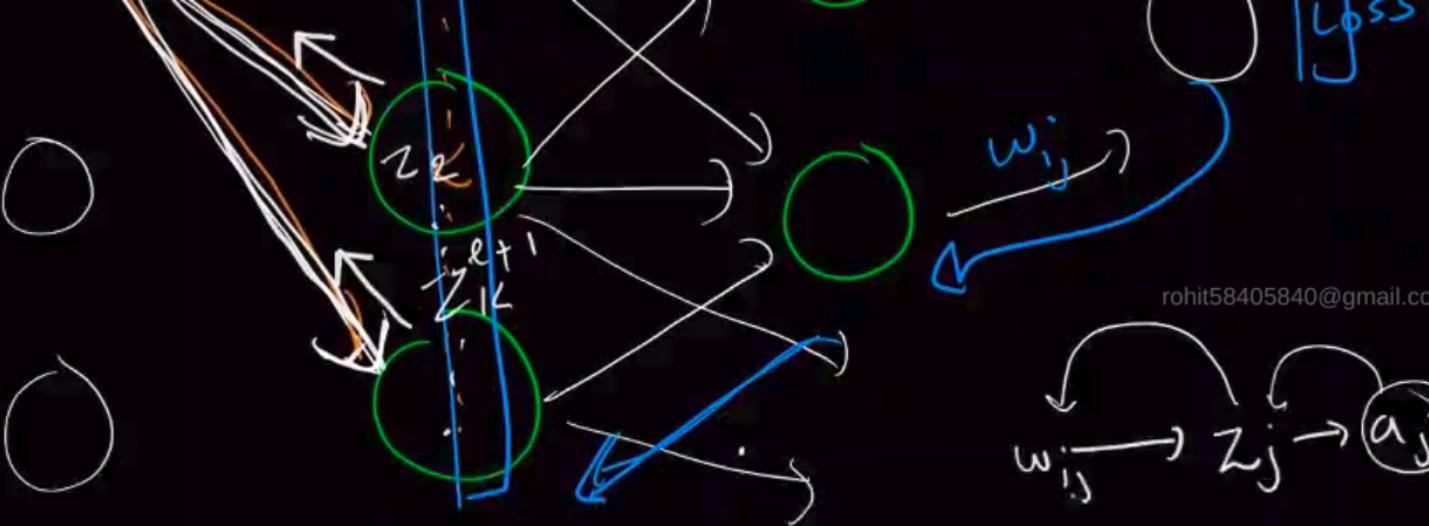




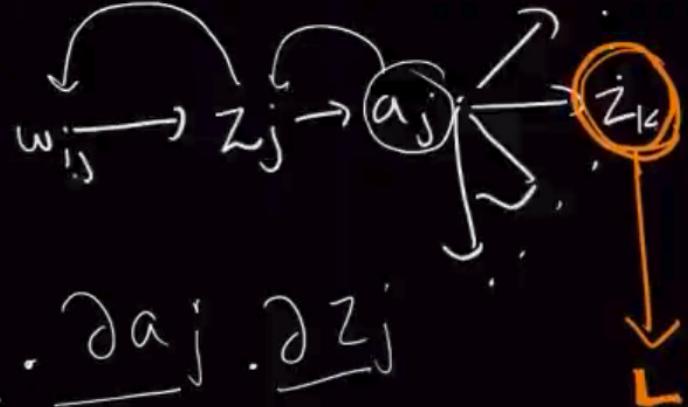
rohit58405840@gmail.com



$$\frac{\partial L}{\partial w_{ij}^l} = \sum_k \frac{\partial L}{\partial z_k^{l+1}}$$



rohit58405840@gmail.com



$$\frac{\partial L}{\partial w_{ij}^l} = \sum_k \frac{\partial L}{\partial z_k^{l+1}} \cdot \frac{\partial z_k^{l+1}}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_{ij}}$$



$$\frac{\partial L}{\partial w_{ij}^l} = \sum_k \frac{\partial L}{\partial z_k^{l+1}} \cdot \frac{\partial z_k^{l+1}}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_{ij}^l}$$
  

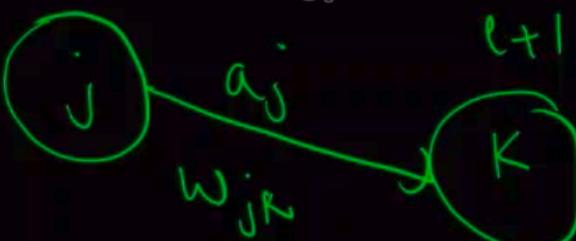
$$= \sum_k S_k^{l+1}$$



$$\frac{\partial L}{\partial w_{ij}^l} = \sum_k \left[ \frac{\partial L}{\partial z_k^{l+1}} \cdot \frac{\partial z_k}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_{ij}^l} \right]$$

$$= \sum_k S_k^{l+1} \cdot w_{jk}$$

rohit58405840@gmail.com



$$z_k^{l+1} = \sum_j w_{jk} \cdot a_j^{l+1}$$

$$\frac{\partial z_k^{l+1}}{\partial a_j^{l+1}} = w_{jk}$$

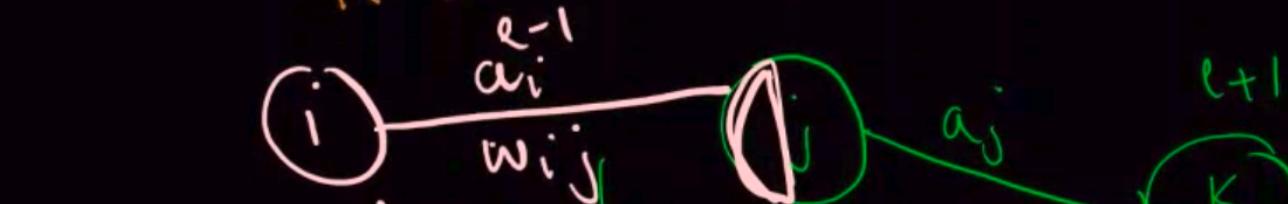
$$\frac{\partial L}{\partial w_{ij}^l} = \sum_k \left( \frac{\partial L}{\partial z_k^{l+1}} \right) \frac{\partial z_k^{l+1}}{\partial a_j^l} \frac{\partial a_j^l}{\partial z_j^l} \frac{\partial z_j^l}{\partial w_{ij}^l}$$



II

$$= \sum_k \left( \sum_{j=1}^{l+1} w_{jk} \right) \sigma'(z_j^l)$$

ohit58405840@gmail.com



$$a_j = \sigma(z_j)$$

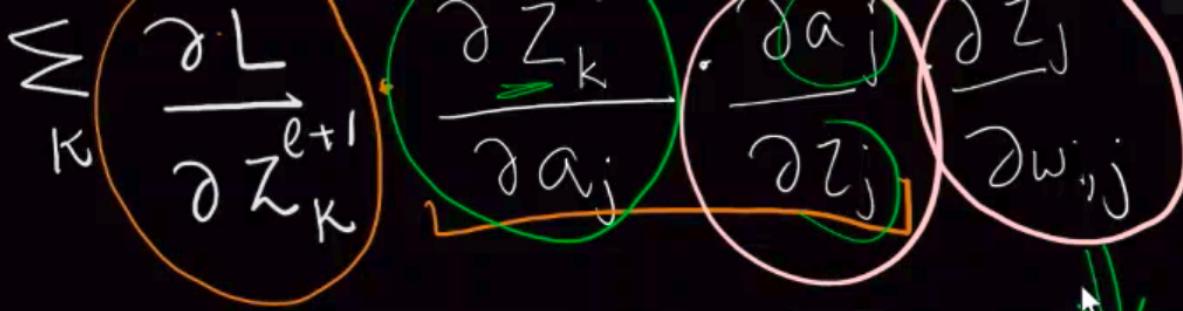
$$\frac{\partial a_j}{\partial z_j} = \sigma'(z_j)$$

$$z_k^{l+1} = \sum_j w_{jk} \cdot a_j^l$$

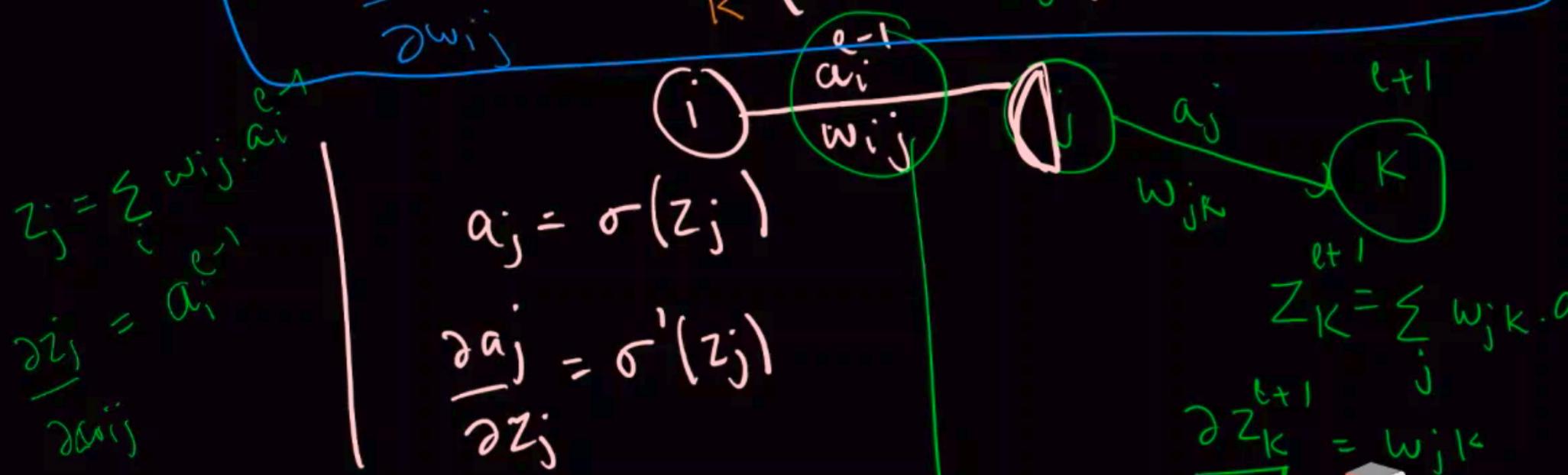
$$\frac{\partial z_k^{l+1}}{\partial a_i^l} = w_{ik}$$



$$\frac{\partial L}{\partial w_{ij}} = \sum_k \left( \frac{\partial L}{\partial z_k^{l+1}} \right) \frac{\partial z_k}{\partial a_j} \frac{\partial a_j}{\partial z_j} \frac{\partial z_j}{\partial w_{ij}}$$



$$\frac{\partial L}{\partial w_{ij}} = \sum_k (S_k^{l+1} \cdot w_{jk}) \sigma'(z_j^l) a_{j-1}^{l-1}$$



$$\frac{\partial z_k^{l+1}}{\partial a_j^l} = w_{jk}$$





$$= \sum_k (S_{K-1}^{l+1} \cdot w_{jk}) \sigma'(z_j^l) a_j^{l-1}$$

$$\frac{\partial L_e}{\partial w_{ij}}$$

$$z_j = \sum_i w_{ij} \cdot a_i^l$$

$$\frac{\partial z_j}{\partial w_{ij}} = a_i^l$$

$$a_j = \sigma(z_j)$$

$$\frac{\partial a_j}{\partial z_j} = \sigma'(z_j)$$

$$a_i^{l-1}$$

i

j

$$a_j$$

$$w_{jk}$$

K

$$a_j^{l+1}$$

$$z_K = \sum_j w_{jk} \cdot a_j^{l+1}$$

$$\frac{\partial z_K}{\partial a_j} = w_{jk}$$

rohit58405840@gmail.com



$$= \sum_{K} (\sum_{k=1}^{l+1} w_{jk}^k) \sigma'(z_j^l) a_k^{l+1}$$

$$\frac{\partial L}{\partial w_{ij}}$$

$$z_j = \sum_i w_{ij} a_i$$

$$a_j = \sigma(z_j)$$

$$\frac{\partial a_j}{\partial z_j} = \sigma'(z_j)$$

$$a_i^{l-1}$$

$$a_j^{l+1}$$

$$w_{jk}$$

$$a_k^{l+1}$$

$$z_K = \sum_j w_{jk} a_j$$

$$\frac{\partial z_K}{\partial a_j} = w_{jk}$$

$$\delta_j^l = \sum_{K} (\sum_{k=1}^{l+1} w_{jk}^k) \sigma'(z_j^l)$$



$$L_e = \sum_k (\delta_k^{l+1} \cdot w_{jk}) \sigma'(z_j^l) a_j^{l+1}$$

$$\frac{\partial L_e}{\partial w_{ij}} = a_i^{l-1} \cdot \delta_j^{l+1} \cdot a_j^{l+1}$$

$$a_j = \sigma(z_j)$$

$$\frac{\partial a_j}{\partial z_j} = \sigma'(z_j)$$

$$\delta_K^{l+1} = \sum_j w_{jk} \cdot a_j^{l+1}$$

$$\frac{\partial \delta_K^{l+1}}{\partial a_j} = w_{jk}$$

$$\delta_j^l = \sum_k (\delta_k^{l+1} \cdot w_{jk}) \sigma'(z_j^l)$$



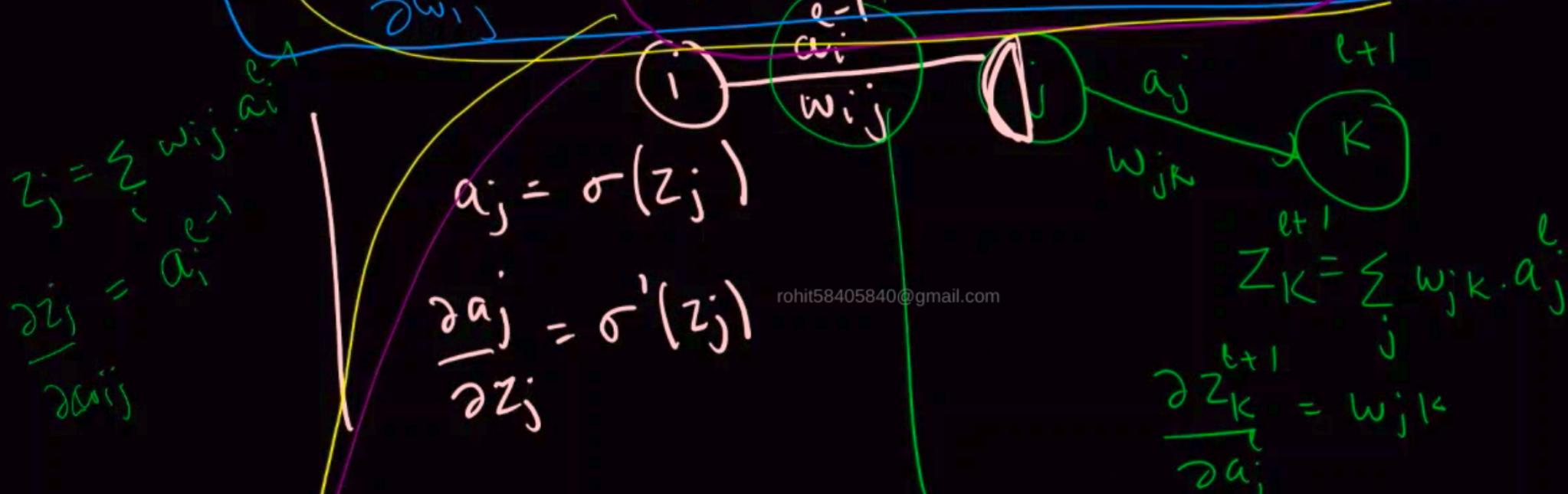
$$z_j = \sum_i w_{ij} \cdot a_i^{l-1}$$

$$\frac{\partial z_j}{\partial w_{ij}} = a_i^{l-1}$$

$$a_j = \sigma(z_j)$$

$$\frac{\partial a_j}{\partial z_j} = \sigma'(z_j)$$

$$= \sum_k (S_K \cdot w_{jk}) \sigma'(z_j^l)$$



rohit58405840@gmail.com

$$Z_K = \sum_j w_{jk} \cdot a_j^l$$

$$\frac{\partial Z_K}{\partial a_j} = w_{jk}$$

$$S_K^l = \sum_j (S_K \cdot w_{jk}) \sigma'(z_j^l)$$

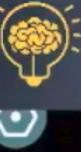




$$\delta_j^l = \sum_k (\delta_k \cdot w_{jk}) \sigma'(z_j^l)$$

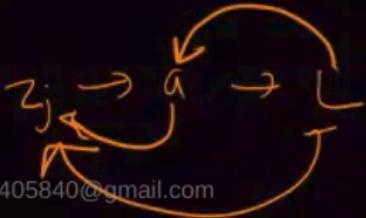
rohit58405840@gmail.com

$$\frac{\partial L}{\partial w_{ij}} = \delta_j^l a_i^{l-1}$$



$$\frac{\partial L}{\partial b_j} = \delta_j^L \cdot 1$$

rohit58405840@gmail.com



For Output layer

where -

$$① \frac{\partial L}{\partial w_{ij}} = \delta_j^L \cdot a_i^{L-1}$$

$$\delta_j^L = (a_j - y_j) \odot \sigma'(z_j)$$

$$② \frac{\partial L}{\partial b_j} = \delta_j^L \cdot 1$$





$$\frac{\partial L}{\partial b_j} = \delta_{j+1} \cdot 1$$

For output layer

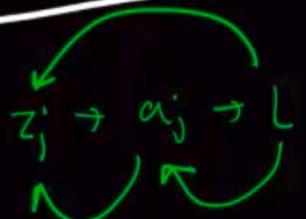
$$\text{1/ } \frac{\partial L}{\partial z_j} = \frac{\partial L}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_j}$$

where -

$$\delta_j^L = (a_j - y_j) \odot \sigma'(z_j)$$

①  $\frac{\partial L}{\partial w_{ij}} = \delta_j^L \cdot a_i^{L-1}$

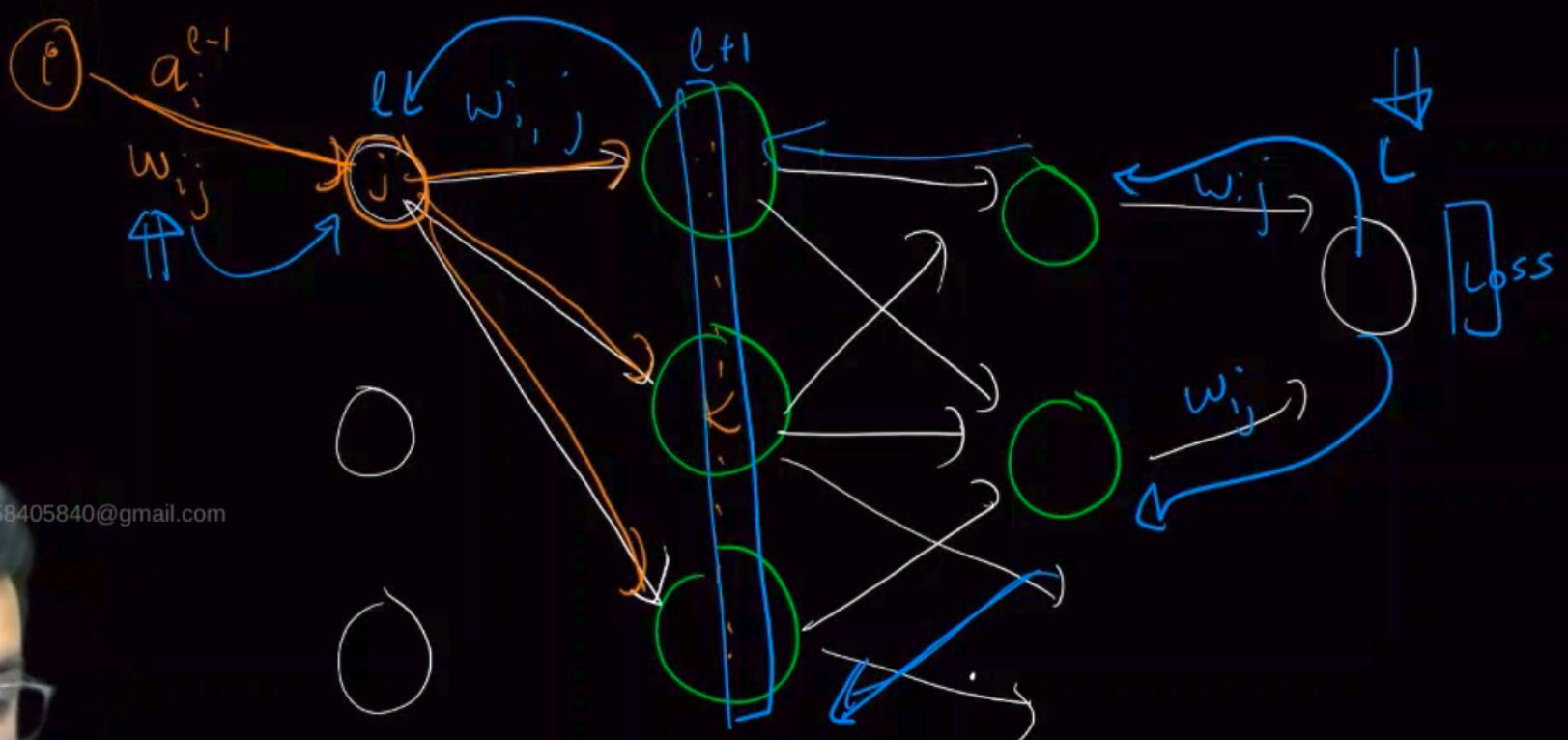
②  $\frac{\partial L}{\partial b_j} = \delta_j^L \cdot 1$



rohit58405840@gmail.com



Coding Blocks Online



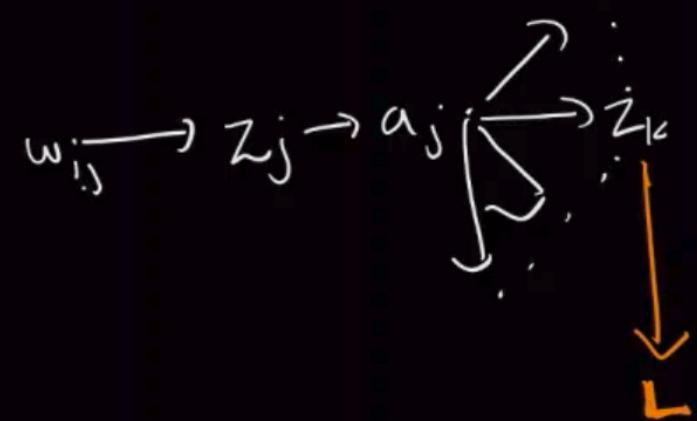
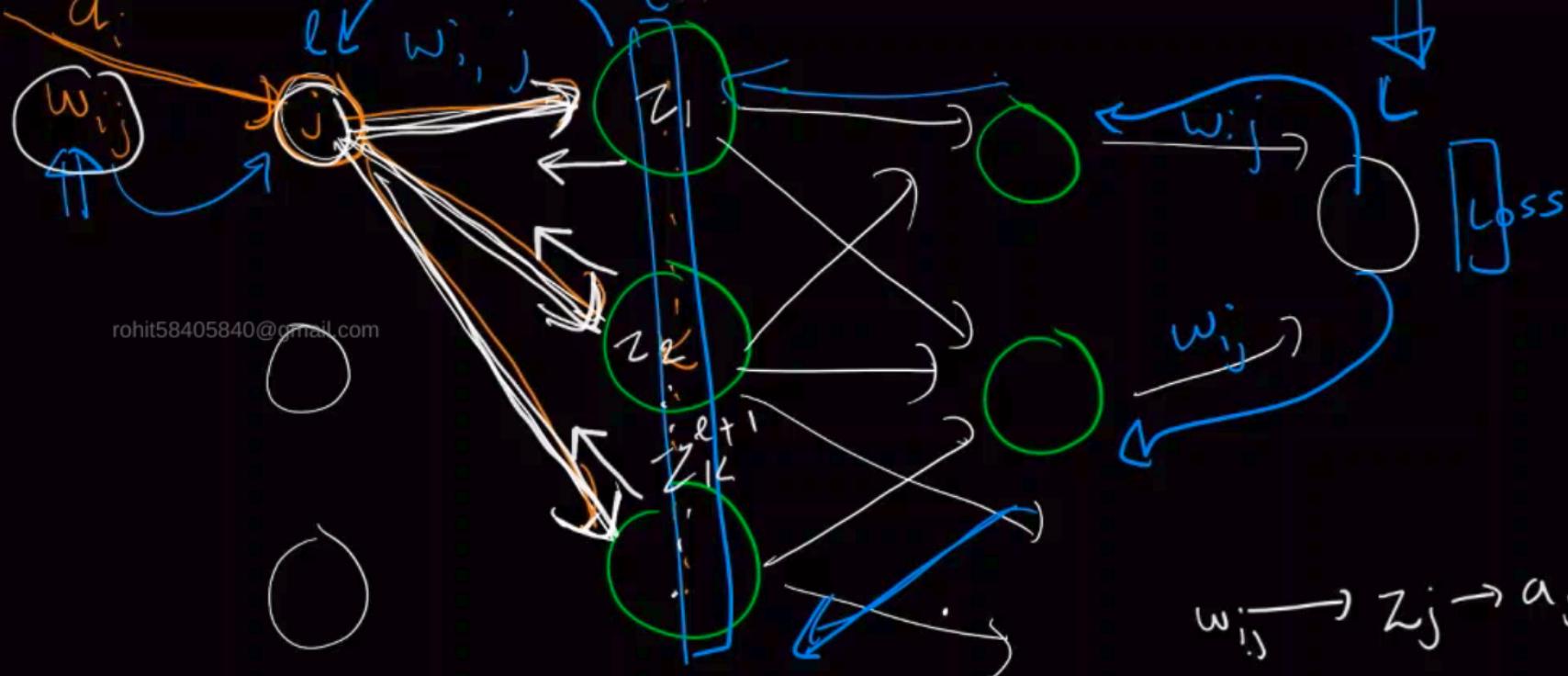
rohit58405840@gmail.com

$$\frac{\partial L}{\partial w_{ij}}$$

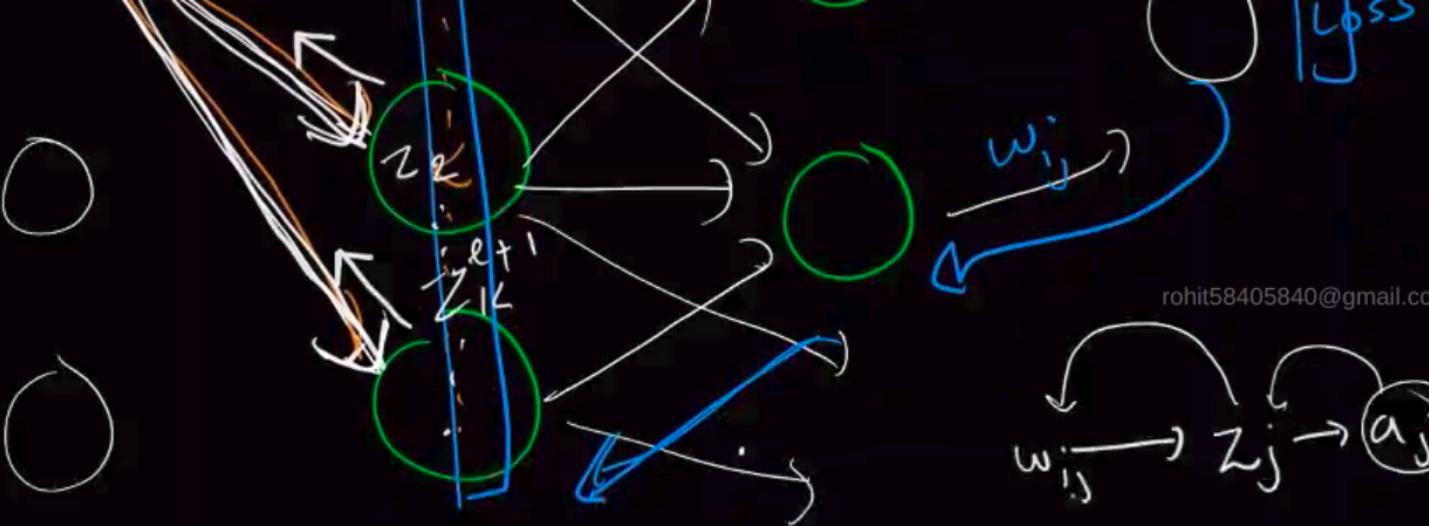




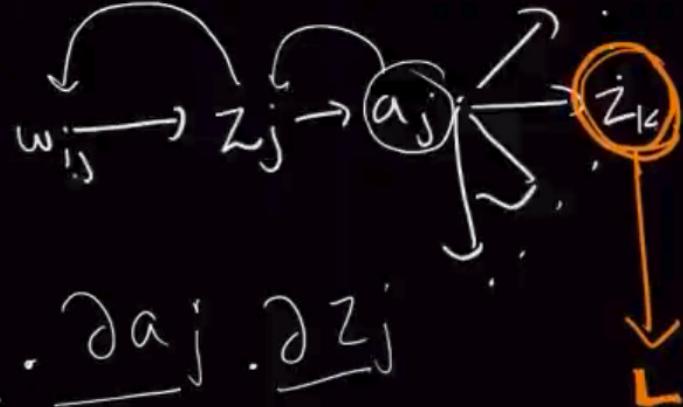
rohit58405840@gmail.com



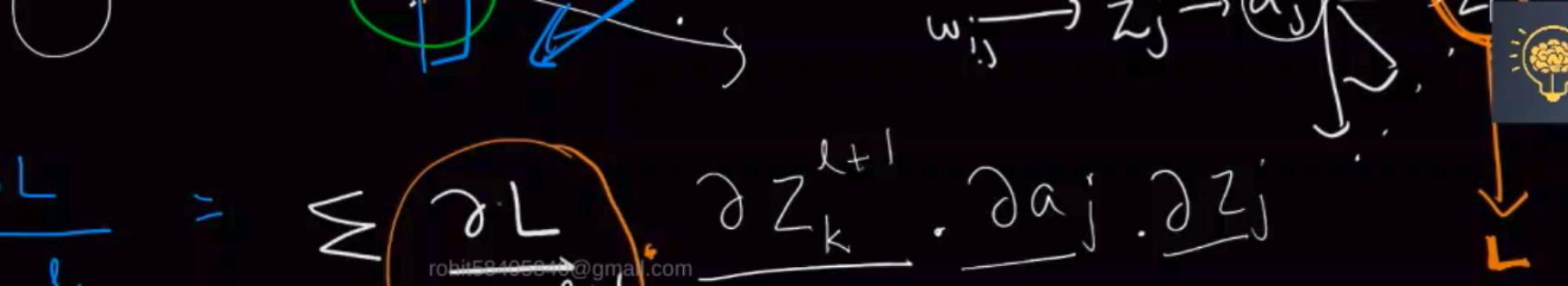
$$\frac{\partial L}{\partial w_{ij}^l} = \sum_k \frac{\partial L}{\partial z_k^{l+1}}$$



rohit58405840@gmail.com



$$\frac{\partial L}{\partial w_{ij}^l} = \sum_k \frac{\partial L}{\partial z_k^{l+1}} \cdot \frac{\partial z_k^{l+1}}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_{ij}}$$



$$\frac{\partial L}{\partial w_{ij}^l} = \sum_k \frac{\partial L}{\partial z_k^{l+1}} \cdot \frac{\partial z_k^{l+1}}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_{ij}^l}$$
  

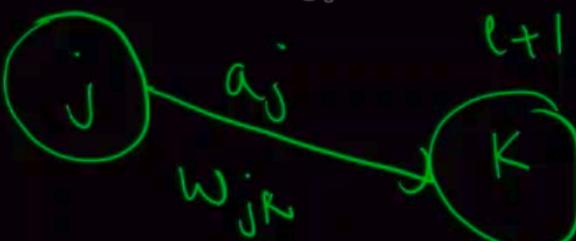
$$= \sum_k S_k^{l+1}$$



$$\frac{\partial L}{\partial w_{ij}^l} = \sum_k \left[ \frac{\partial L}{\partial z_k^{l+1}} \cdot \frac{\partial z_k}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_{ij}^l} \right]$$

$$= \sum_k S_k^{l+1} \cdot w_{jk}$$

rohit58405840@gmail.com



$$z_k^{l+1} = \sum_j w_{jk} \cdot a_j^{l+1}$$

$$\frac{\partial z_k^{l+1}}{\partial a_j^{l+1}} = w_{jk}$$

$$\frac{\partial L}{\partial w_{ij}^l} = \sum_k \left( \frac{\partial L}{\partial z_k^{l+1}} \right) \frac{\partial z_k^{l+1}}{\partial a_j^l} \frac{\partial a_j^l}{\partial z_j^l} \frac{\partial z_j^l}{\partial w_{ij}^l}$$



II

$$= \sum_k \left( \sum_{j=1}^{l+1} w_{jk} \right) \sigma'(z_j^l)$$

ohit58405840@gmail.com



$$a_j = \sigma(z_j)$$

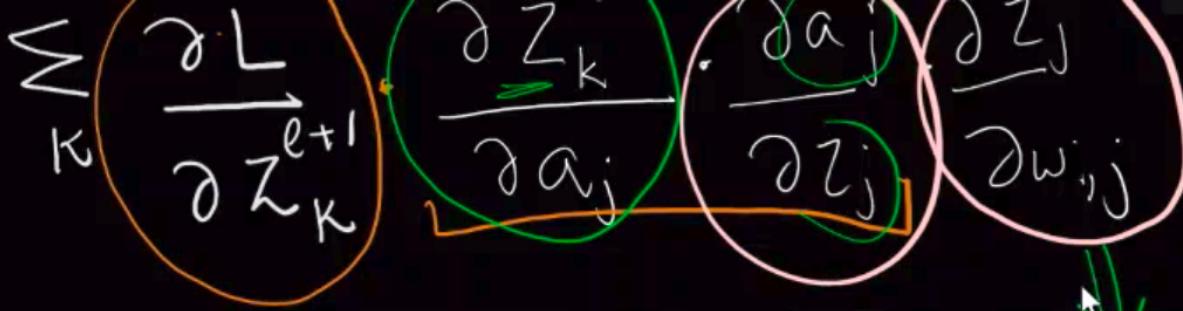
$$\frac{\partial a_j}{\partial z_j} = \sigma'(z_j)$$

$$z_k^{l+1} = \sum_j w_{jk} \cdot a_j^l$$

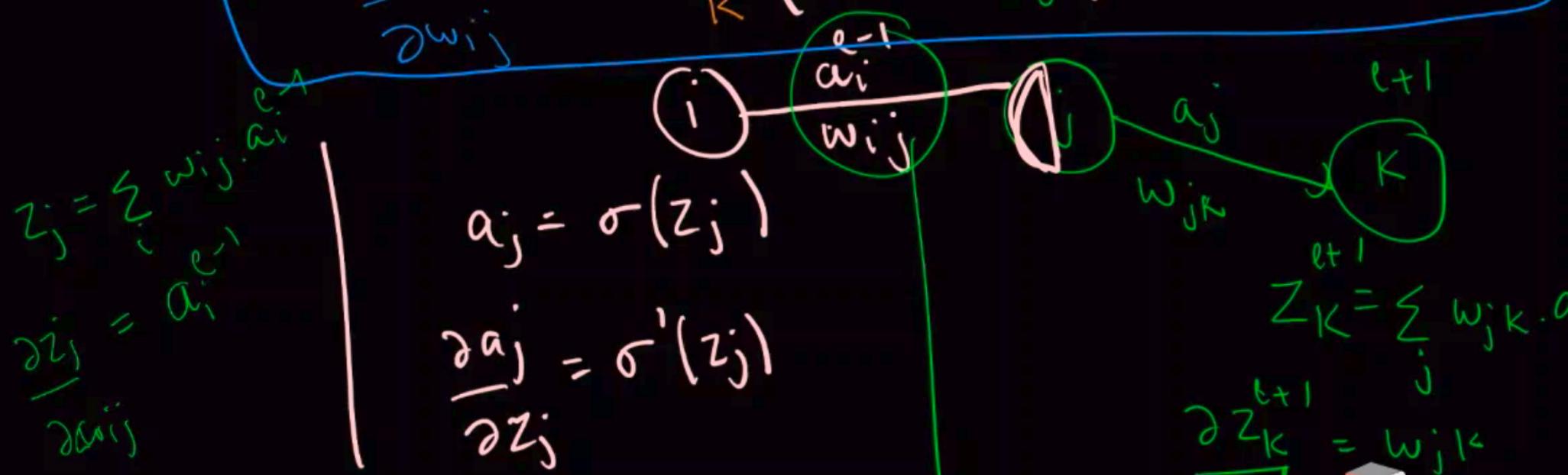
$$\frac{\partial z_k^{l+1}}{\partial a_i^l} = w_{ik}$$



$$\frac{\partial L}{\partial w_{ij}} = \sum_k \left( \frac{\partial L}{\partial z_k^{l+1}} \right) \frac{\partial z_k}{\partial a_j} \frac{\partial a_j}{\partial z_j} \frac{\partial z_j}{\partial w_{ij}}$$



$$\frac{\partial L}{\partial w_{ij}} = \sum_k (S_k^{l+1} \cdot w_{jk}) \sigma'(z_j^l) a_{j-1}^{l-1}$$



$$\frac{\partial z_k^{l+1}}{\partial a_j^l} = w_{jk}^{l+1}$$





$$= \sum_k (S_K^{l+1} \cdot w_{jk}) \sigma'(z_j^l) a_i^{l-1}$$

$$\frac{\partial L_e}{\partial w_{ij}}$$

$$z_j = \sum_i w_{ij} \cdot a_i^l$$

$$\frac{\partial z_j}{\partial w_{ij}} = a_i^l$$

$$a_j = \sigma(z_j)$$

$$\frac{\partial a_j}{\partial z_j} = \sigma'(z_j)$$

$$a_i^{l-1}$$

$$a_j$$

$$w_{jk}$$

$$a_j^{l+1}$$

$$z_k = \sum_j w_{jk} \cdot a_j^l$$

$$\frac{\partial z_k}{\partial a_j} = w_{jk}$$

rohit58405840@gmail.com



$$= \sum_{K} (\delta_{K}^{l+1} \cdot w_{j|K}) \sigma'(z_j^l) a_j^{l+1}$$

$$\frac{\partial L}{\partial w_{ij}}$$

$$z_j = \sum_i w_{ij} a_i$$

$$a_j = \sigma(z_j)$$

$$\frac{\partial a_j}{\partial z_j} = \sigma'(z_j)$$

$$\delta_{j|K}^{l+1}$$

$$a_j^{l+1}$$

$$w_{j|K}$$

$$z_K = \sum_j w_{j|K} a_j^l$$

$$\frac{\partial z_K}{\partial a_j^l} = w_{j|K}$$

$$\delta_j^l = \sum_K (\delta_{K}^{l+1} \cdot w_{j|K}) \sigma'(z_j^l)$$



$$L_e = \sum_k (\delta_k^{l+1} \cdot w_{jk}) \sigma'(z_j^l) a_j^{l+1}$$

$$\frac{\partial L_e}{\partial w_{ij}} = a_i^{l-1} \cdot \delta_j^{l+1} \cdot a_j^{l+1}$$

$$z_j = \sum_i w_{ij} \cdot a_i^{l-1}$$

$$\frac{\partial z_j}{\partial w_{ij}} = a_i^{l-1}$$

$$a_j = \sigma(z_j)$$

$$\frac{\partial a_j}{\partial z_j} = \sigma'(z_j)$$

$$z_K = \sum_j w_{jk} \cdot a_j^{l+1}$$

$$\frac{\partial z_K}{\partial a_j} = w_{jk}$$

$$\delta_j^l = \sum_k (\delta_k^{l+1} \cdot w_{jk}) \sigma'(z_j^l)$$



$$\frac{\partial L_e}{\partial w_{ij}} = \sum_k \left( S_K^{l+1} \cdot w_{jik} \right) \sigma'(z_j^e) \cdot a_i^{e-1}$$
$$z_j = \sum_i w_{ij} \cdot a_i^{e-1}$$
$$\frac{\partial z_j}{\partial w_{ij}} = a_i^{e-1}$$
$$a_j = \sigma(z_j)$$
$$\frac{\partial a_j}{\partial z_j} = \sigma'(z_j)$$
$$S_K^{l+1} = \sum_j w_{jk} \cdot a_j^e$$
$$\frac{\partial z_k}{\partial a_j} = w_{jk}$$

$$S_K^e = \sum_k \left( S_K \cdot w_{jik} \right) \sigma'(z_j^e)$$





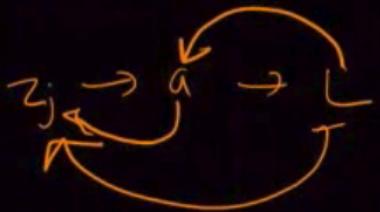
$$\delta_j^l = \sum_k (\delta_k \cdot w_{jk}) \sigma'(z_j^l)$$

rohit58405840@gmail.com

$$\frac{\partial L}{\partial w_{ij}} = \delta_j^l a_i^{l-1}$$



$$\frac{\partial L}{\partial b_j} = \delta_j^L \cdot 1$$



For output layer

$$1, \frac{\partial L}{\partial z_j} = \frac{\partial L}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_j}$$

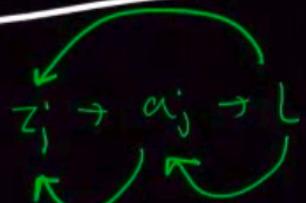
where -

$$\delta_j^L = (a_j - y_j) \odot \sigma'(z_j)$$

$$\textcircled{1} \quad \frac{\partial L}{\partial w_{ij}} = \delta_j^L \cdot a_i^{L-1}$$

$$\textcircled{2} \quad \frac{\partial L}{\partial b_j} = \delta_j^L \cdot 1$$

rohit58405840@gmail.com



$$\delta_j = \sum_k (\omega_{jk} \cdot \delta_k) \sigma'(z_j)$$



Hidden Neurons

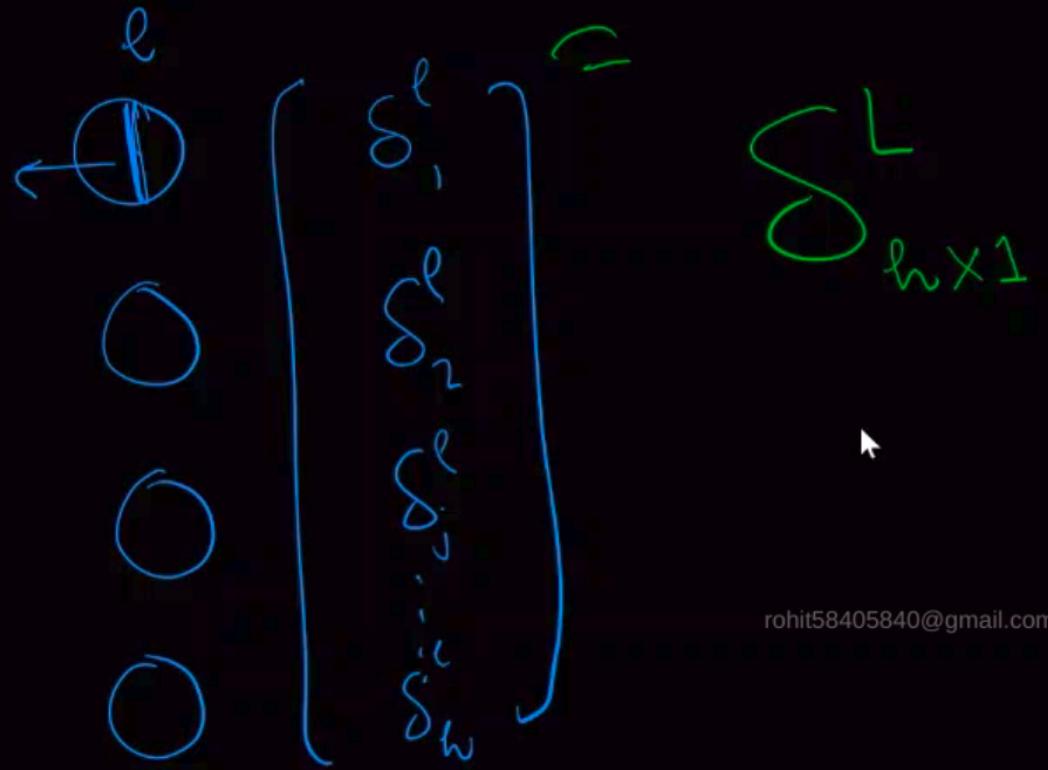
(1)

$$\frac{\partial L}{\partial w_{ij}} = \delta_j^l a_i^{l-1}$$

(2)

$$\frac{\partial L}{\partial b} = \delta_j^l$$

$$\delta_j^l = \sum_k (\omega_{jk}^{l+1} \cdot \delta_k^{l+1}) \sigma'(z_j^l)$$



rohit58405840@gmail.com



$$\frac{\partial L}{\partial w_{i,j}}$$

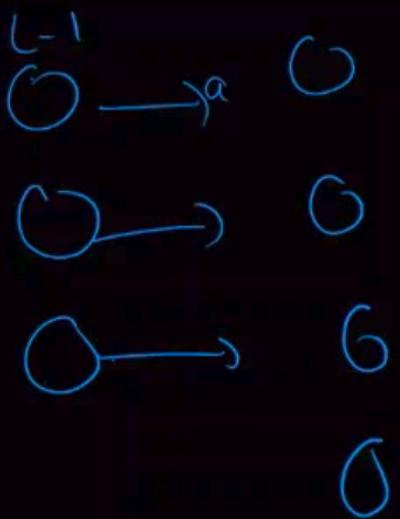
$$\frac{\partial L}{\partial w^l} \rightarrow a^{l-1} \cdot (\varepsilon^l)^T$$



$$\frac{\partial L}{\partial \underline{w^l}} = a^{l-1} \cdot (\underline{\delta^l})^T$$

rohit58405840@gmail.com

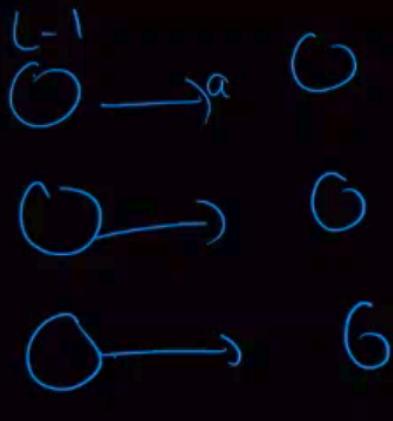
$$w^- \begin{bmatrix} dw & dw & \dots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix}_{(l-1, l)}$$





$$W^T = \begin{bmatrix} dw & dw & \dots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \end{bmatrix}_{(l-1, l)}$$

rohit58405840@gmail.com



~~Corrector~~

$$\left[ \begin{array}{c} a_1^{l-1} \\ a_2^{l-1} \\ a_3^{l-1} \\ \vdots \end{array} \right] \left[ \begin{array}{c} \delta_1 \\ a_2^{l-2} \\ a_3^{l-3} \\ \vdots \end{array} \right]$$

