

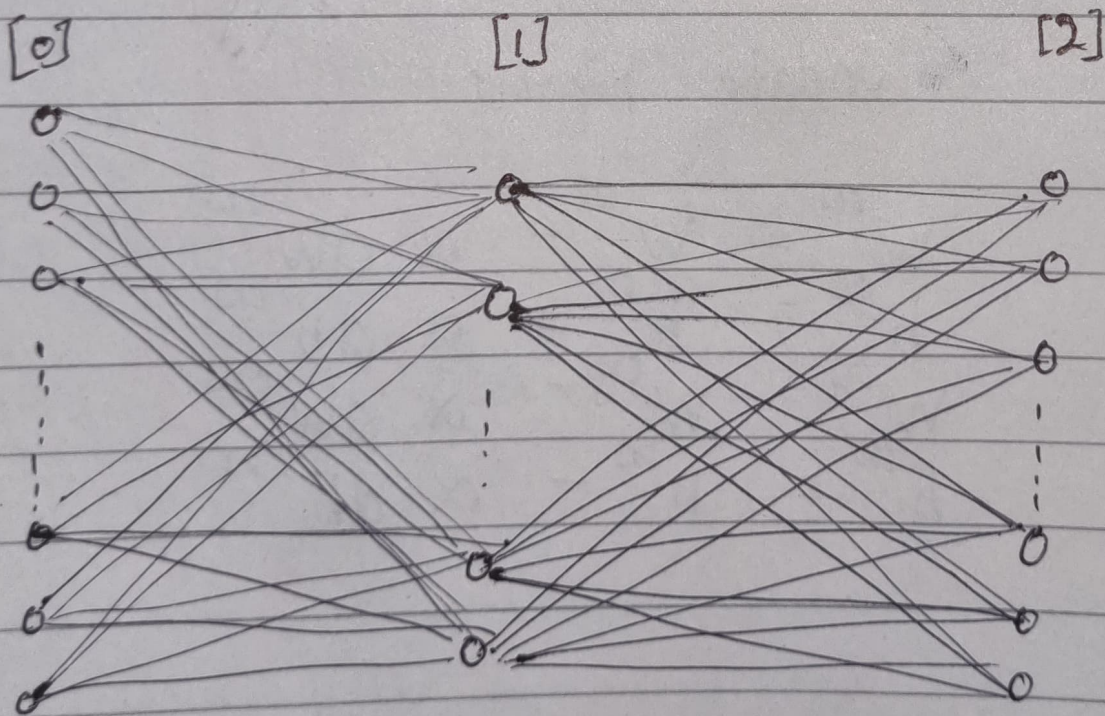
image = 784

m = no of image

$$X = \begin{bmatrix} \text{---} X_1 \text{---} \\ \text{---} X_2 \text{---} \\ \text{---} \vdots \text{---} \\ \text{---} X_m \text{---} \end{bmatrix}_m$$

Transpose to work easily

$$X^T = \begin{bmatrix} | & | & \dots & | \\ X_1 & X_2 & \dots & X_m \\ | & | & \dots & | \end{bmatrix}_{784}$$



input layer
784

hidden layer

output layer
10

back prop

forward prop

$$Y = \text{one hot}(Y) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}^T$$

labels ~~into~~ encoded

$$dz^{[2]} = A^{[2]} - Y$$

$10 \times m$ $10 \times m$ $10 \times m$

$$dW^{[2]} = \frac{1}{m} dz^{[2]} \cdot A^{[1]T}$$

$10 \times \text{hdL}$ $10 \times m$ $m \times \text{hdL}$

$$db^{[2]} = \frac{1}{m} \sum dz^{[2]}$$

10×1 10×1

$$dz^{[1]} = W^{[2]T} \cdot dz^{[2]} * \text{ReLU}'(z^{[1]})$$

$\text{hdL} \times m$ $\text{hdL} \times 10$ $10 \times m$ $\text{hdL} \times m$

$$dW^{[1]} = \frac{1}{m} dz^{[1]} \cdot A^{[0]T}$$

$\text{hdL} \times 784$ $\text{hdL} \times m$ $m \times 784$

$$db^{[1]} = \frac{1}{m} \sum dz^{[1]}$$

$\text{hdL} \times 1$ $\text{hdL} \times 1$

$$A^{[0]} = X$$

$$Z^{[1]} = W^{[1]} \cdot A^{[0]} + b^{[1]}$$

$\text{hdL} \times m$ $\text{hdL} \times 784$ $784 \times m$ $\text{hdL} \times 1$

$$A^{[1]} = \text{ReLU}(Z^{[1]})$$

$$\text{ReLU}(x) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

$$Z^{[2]} = W^{[2]} \cdot A^{[1]} + b^{[2]}$$

$10 \times m$ $10 \times \text{hdL}$ $\text{hdL} \times m$ 10×1

$$A^{[2]} = \text{softmax}(Z^{[2]})$$

$$\hat{y}(z)_i = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}}$$

update params

$$\begin{aligned} W^{[1]} &= W^{[1]} - \alpha dW^{[1]} \\ b^{[1]} &= b^{[1]} - \alpha db^{[1]} \\ W^{[2]} &= W^{[2]} - \alpha dW^{[2]} \\ b^{[2]} &= b^{[2]} - \alpha db^{[2]} \end{aligned}$$