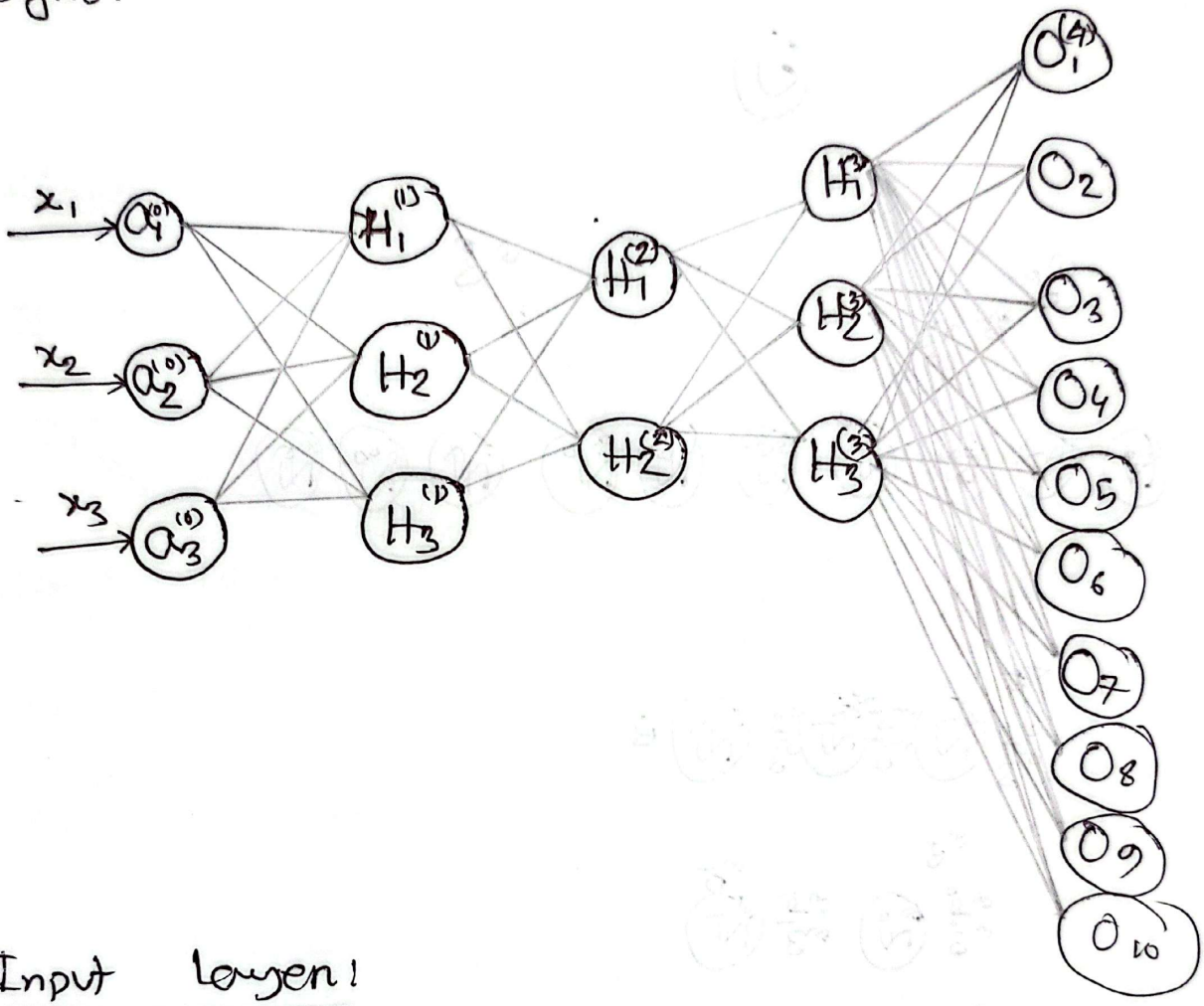
Md. Rayhanul Islam
2010976154

Q: Draw a neural network having 3 hidden layers for classifying images of 10 English digits.



Input layer:

$$z_i^1 = w_i a_i^0 + b_i^1 \quad \text{where } i = 0, 1, 2$$

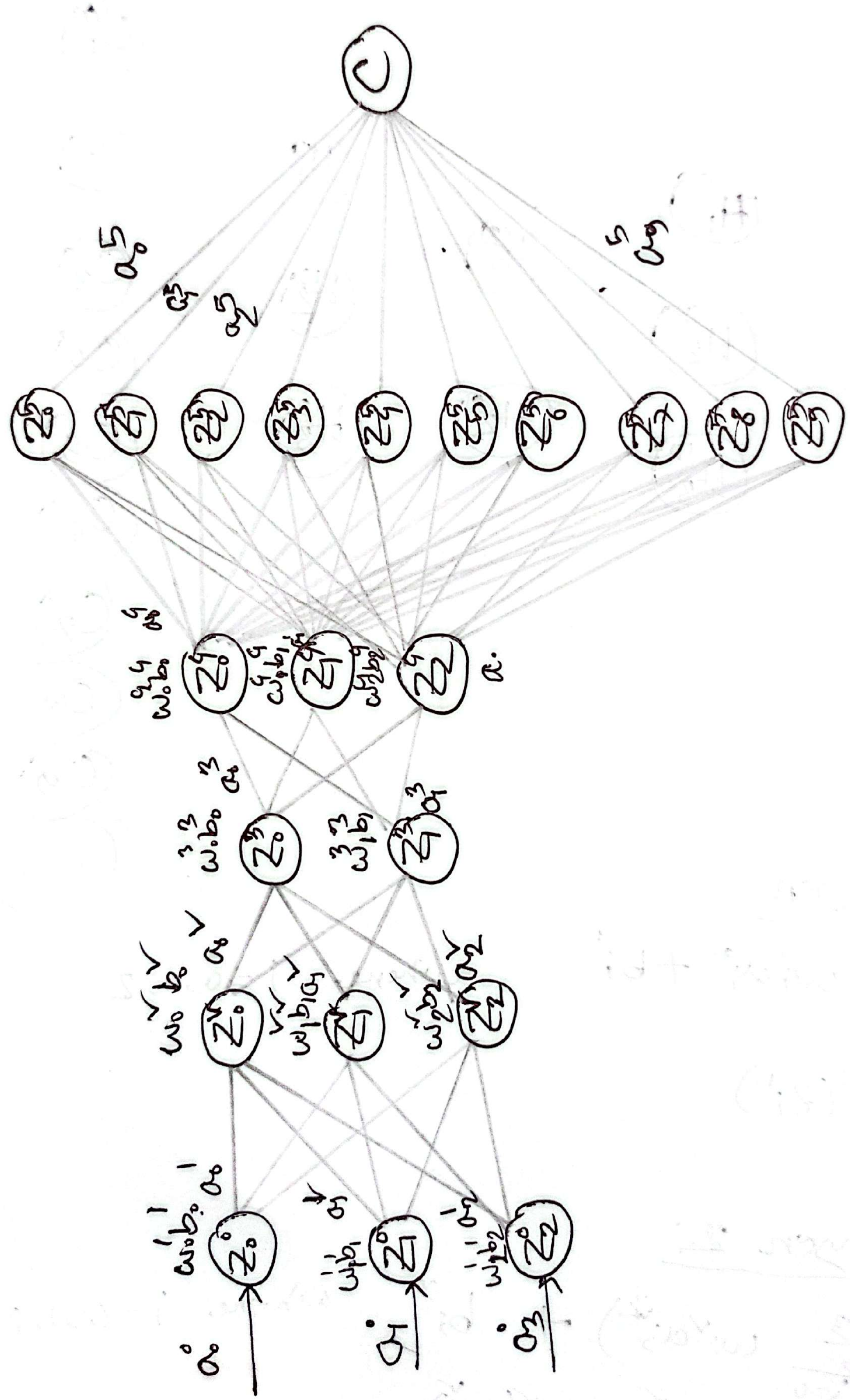
$$a_i^1 = g^1(z_i^1)$$

Hidden layer 2:

$$z_i^2 = \left(\sum_{j=0}^{3} w_i^2 a_i^1\right) + b_i^2 \quad \text{where } i = 0, 1, 2$$

$$a_i^2 = g^2(z_i^2)$$

Forward Propagation:



Hidden layer 3:

$$z_i^3 = \left(\sum_{j=0}^{2} w_{ij}^3 a_j^2\right) + b_i^3 \quad [i = 0,1]$$

$$a_i^3 = g^3(z_i^3)$$

Hidden layer 4:

$$z_i^4 = \sum_{j=0}^{i} w_{ij}^4 a_j^3 + b_i^4 \quad [i = 0,1,2]$$

$$a_i^4 = g^4(z_i^4)$$

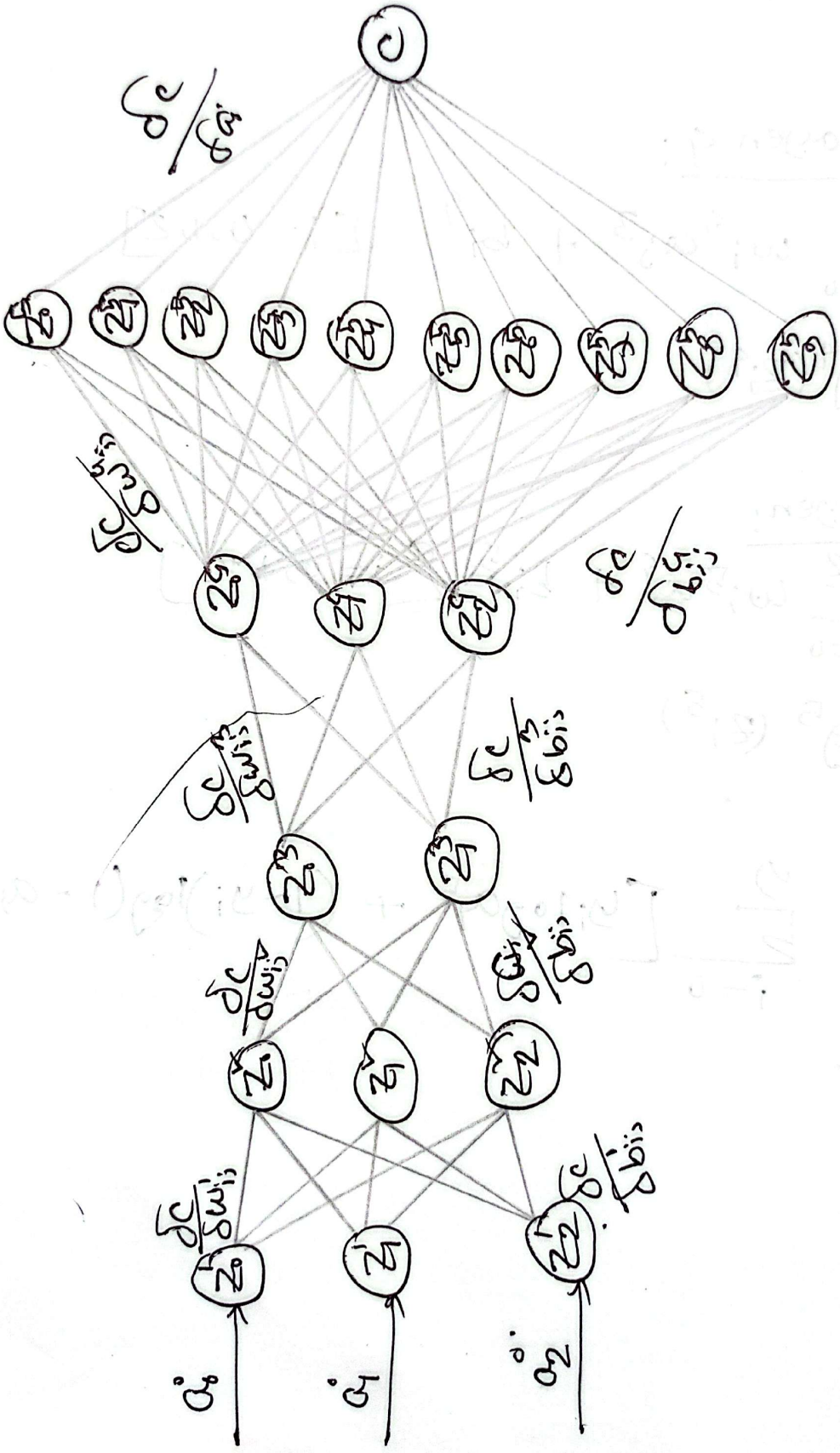
Output layer:

$$z_i^5 = \sum_{j=0}^{2} w_{ij}^5 a_j^4 + b_i^5 \quad [i = 0,1,2]$$

$$a_j^5 = g^5(z_i^5)$$

$$\therefore C = \sum_{i=0}^{2} \left[y_i \log a_i^5 + (1-y_i)\log(1-a_i^5)\right]$$

Backpropogation:



Hidden layer 4:

$$\frac{\partial C}{\partial w_{ij}^4} = \frac{\partial C}{\partial a_j^4} \cdot \frac{\partial a_j^4}{\partial z_j^4} \cdot \frac{\partial z_j^4}{\partial w_{ij}^4}$$

$$\frac{\partial C}{\partial b_{ij}^4} = \frac{\partial C}{\partial a_j^4} \cdot \frac{\partial a_j^4}{\partial z_j^4} \cdot \frac{\partial z_j^4}{\partial b_{ij}^4}$$

Hidden layer 3:

$$\frac{\partial C}{\partial w_{ij}^3} = \frac{\partial C}{\partial a_j^4} \cdot \frac{\partial a_j^4}{\partial z_j^4} \cdot \frac{\partial z_j^4}{\partial a_j^3} \cdot \frac{\partial a_j^3}{\partial z_j^3} \cdot \frac{\partial z_j^3}{\partial w_{ij}^3}$$

$$\frac{\partial C}{\partial b_{ij}^3} = \frac{\partial C}{\partial a_j^4} \cdot \frac{\partial a_j^4}{\partial z_j^4} \cdot \frac{\partial z_j^4}{\partial a_j^3} \cdot \frac{\partial a_j^3}{\partial z_j^3} \cdot \frac{\partial z_j^3}{\partial b_{ij}^3}$$

Hidden layer 2:

$$\frac{\partial C}{\partial w_{ij}^2} = \frac{\partial C}{\partial a_j^4} \cdot \frac{\partial a_j^4}{\partial z_j^4} \cdot \frac{\partial z_j^4}{\partial a_j^3} \cdot \frac{\partial a_j^3}{\partial z_j^3} \cdot \frac{\partial z_j^3}{\partial a_j^2} \cdot \frac{\partial a_j^2}{\partial z_j^2} \cdot \frac{\partial z_j^2}{\partial w_{ij}^2}$$

$$\frac{\partial C}{\partial b_{ij}^2} = \frac{\partial C}{\partial a_j^4} \cdot \frac{\partial a_j^4}{\partial z_j^4} \cdot \frac{\partial z_j^4}{\partial a_j^3} \cdot \frac{\partial a_j^3}{\partial z_j^3} \cdot \frac{\partial z_j^3}{\partial a_j^2} \cdot \frac{\partial a_j^2}{\partial z_j^2} \cdot \frac{\partial z_j^2}{\partial b_{ij}^2}$$

Layer 1:

$$\frac{\partial C}{\partial w_{ij}^1} = \frac{\partial C}{\partial a_{ij}^4} \cdot \frac{\partial a_j^4}{\partial z_j^4} \cdot \frac{\partial z_j^4}{\partial a_j^3} \cdot \frac{\partial a_j^3}{\partial z_j^3} \cdot \frac{\partial z_j^3}{\partial a_j^2} \cdot \frac{\partial a_j^2}{\partial z_j^2} \cdot \frac{\partial z_j^2}{\partial a_j^1} \cdot \frac{\partial a_j^1}{\partial z_j^1} \cdot \frac{\partial z_j^1}{\partial w_{ij}^1}$$

$$\frac{\partial C}{\partial b_{ij}^1} = \frac{\partial C}{\partial a_{ij}^4} \cdot \frac{\partial a_{ij}^4}{\partial z_j^4} \cdot \frac{\partial z_j^4}{\partial a_j^3} \cdot \frac{\partial a_j^3}{\partial z_j^3} \cdot \frac{\partial z_j^3}{\partial a_j^2} \cdot \frac{\partial a_j^2}{\partial z_j^2} \cdot \frac{\partial z_j^2}{\partial a_j^1} \cdot \frac{\partial a_j^1}{\partial z_j^1} \cdot \frac{\partial z_j^1}{\partial b_{ij}^1}$$