Homework Assignment 4

Problem 1:

$A_1 = (2,10)$, $A_2 = (2,5)$, $A_3 = (8,4)$, $B_1 = (5,8)$,

$C_1 = (1,2)$, $C_2 = (4,9)$

clusters $\rightarrow A_1, B_1, C_2$

For A_1,

$$d(A_2) = \sqrt{(2-2)^2+(10-5)^2} = \sqrt{5^2} = 5$$

$$d(A_3) = \sqrt{(2-8)^2+(10-4)^2} = \sqrt{6^2+6^2} = \sqrt{72} = 8.48$$

$$d(C_1) = \sqrt{(2-1)^2+(10-2)^2} = \sqrt{1+8^2} = 8.06$$

For B_1,

$$d(A_2) = \sqrt{(5-2)^2+(8-5)^2} = \sqrt{3^2+3^2} = 4.24$$

$$d(A_3) = \sqrt{(5-8)^2+(8-4)^2} = \sqrt{3^2+4^2} = 5$$

$$d(C_1) = \sqrt{(5-1)^2+(8-2)^2} = \sqrt{4^2+6^2} = 7.21$$

For C_2,

$$d(A_1) = \sqrt{(4-2)^2+(9-5)^2} = \sqrt{2^2+4^2}$$
$$= 4.47$$

$$d(A_2) = \sqrt{(4-8)^2+(9-4)^2}$$
$$= \sqrt{4^2+5^2} = \sqrt{41} = 6.4$$

$$d(C_1) = \sqrt{(4-1)^2+(9-2)^2} = \sqrt{3^2+7^2}$$
$$= 7.61$$

Using Euclidean distance, we see that the clusters formed are

$A_1 \rightarrow A_1$
$B_1 \rightarrow B_1, A_2, A_3, C_1$
$C_2 \rightarrow C_2$

For centroid calculation,

A_1 & C_2 remain unchanged

For B_1,

$$X = \left(\frac{2+8+5+1}{4}\right) = 4$$

$$Y = \left(\frac{5+4+8+2}{4}\right) = 4.75$$

centroid $\rightarrow$ $(4, 4.75)$

Problem 3:

Gender = M , Car Type = Family , Shirt size = Large

$$P(C_0) = \frac{10}{20} \qquad P(C_1) = \frac{10}{20}$$

$$P(X|C_0) = P(\text{Gender} = M|C_0) \times P(\text{Car Type} = \text{Family}|C_0) \times P(\text{Shirt Size} = \text{Large}|C_0)$$

$$= \frac{6}{10} \times \frac{1}{10} \times \frac{2}{10}$$

$$= \frac{12}{1000}$$

$$P(X|C_1) = P(\text{Gender} = M|C_1) \times P(\text{Car Type} = \text{Family}|C_1) \times P(\text{Shirt size} = \text{Large}|C_1)$$

$$= \frac{4}{10} \times \frac{3}{10} \times \frac{2}{10}$$

$$= \frac{24}{1000}$$

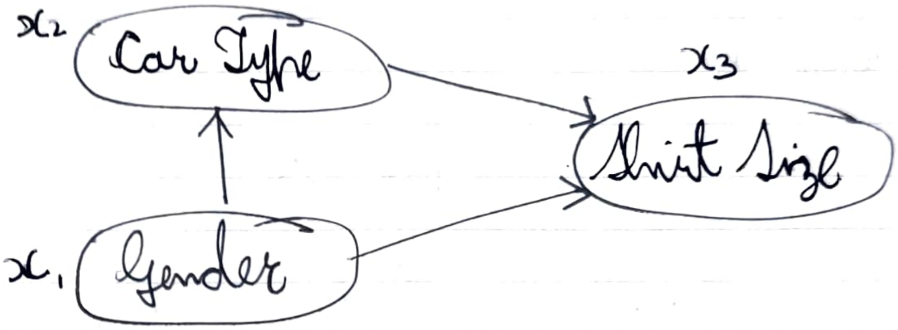
$$P(X|C_0) \times P(C_0) = \frac{12}{100} \times \frac{1}{2}$$

$$= 0.006$$

$$P(X|C_1) \times P(C_1) = \frac{24}{1000} \times \frac{1}{2} = 0.012$$

Since $P(X|C_1)P(C_1) > P(X|C_0)P(C_0)$, the new test example is classified as C_1.

2.



$$P(C|x_1, x_2, x_3) = \frac{P(x_1, x_2, x_3|C)\, P(C)}{P(x_1, x_2, x_3)}$$

$$= \frac{P(x_1|C)\, P(x_2|C, x_1)\, P(x_3|C, x_1, x_2)\, P(C)}{P(x_1, x_2, x_3)}$$

$$P(x_1 = M, x_2 = \text{Family}, x_3 = \text{Large}|C_0)\, P(C_0)$$

$$= \left[\frac{6}{10} \times \frac{1}{6} \times 0\right] \times \frac{1}{2} = 0$$

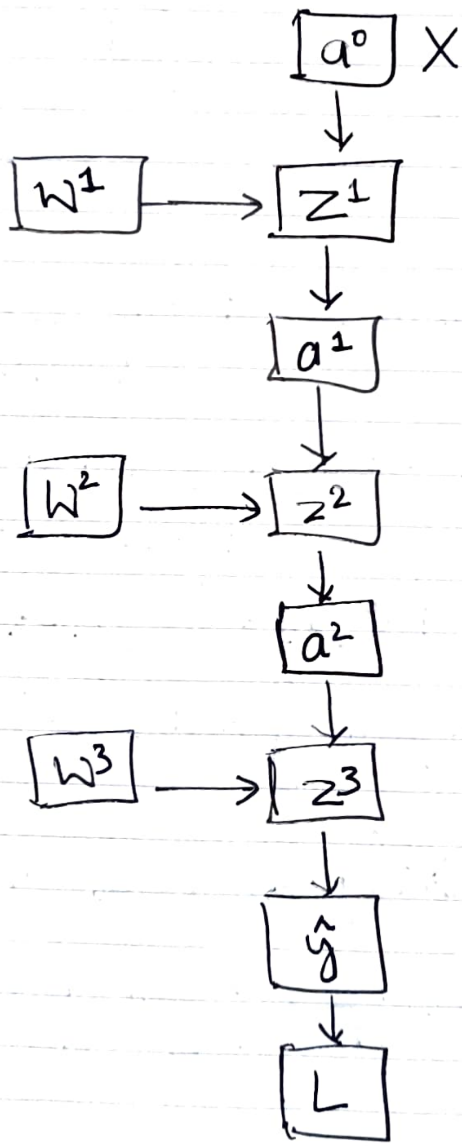
$$P(x_1 = M, x_2 = \text{Family}, x_3 = \text{Large}|C_1)\, P(C_1)$$

$$= \left[\frac{4}{10}, \frac{3}{4}, \frac{1}{10}\right] \times \frac{1}{2} = 0.06$$

$\Rightarrow C_1$

Problem 4:

Forward Propagation
Computation Graph



$$z^1 = W^1 X$$

$$a^1 = \sigma(z^1)$$

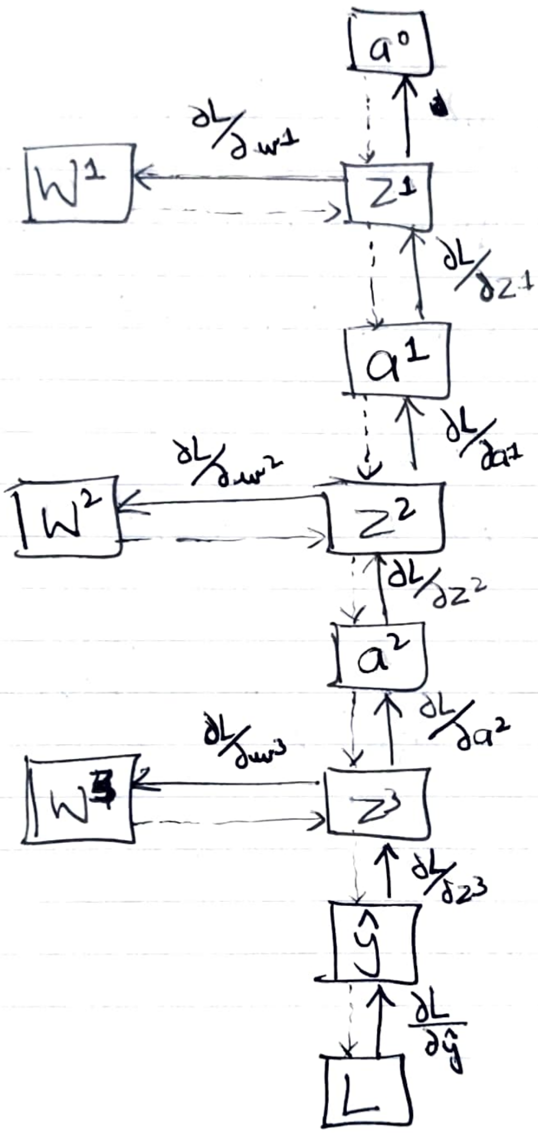
$$z^2 = W^2 a^1$$

$$a^2 = \sigma(z^2)$$

$$z^3 = W^3 a^2$$

$$\hat{y} = \frac{e^{z_i^3}}{\sum_j e^{z_j^3}}$$

Backward Propogation Computation Graph:



$$\frac{\partial L}{\partial w^3} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial z^3} \frac{\partial z^3}{\partial w^3}$$

$$\frac{\partial L}{\partial \hat{y}} = \frac{\partial}{\partial \hat{y}} \left[(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2 \right]$$

$$= 2(y_1 - \hat{y}_1)(-1) + 2(y_2 - \hat{y}_2)(-1)$$

$$= -2(y_1 - \hat{y}_1) - 2(y_2 - \hat{y}_2)$$

$$\frac{\partial \hat{y}}{\partial z^3} = \frac{\partial}{\partial z^3}\left(\frac{e^{z_1}}{\sum_j e^{z_j}}\right) + \frac{\partial}{\partial z^3}\left(\frac{e^{z_2}}{\sum_j e^{z_j}}\right)$$

$$\frac{\partial}{\partial z^3}\left(\frac{e^{z_i}}{\sum_j e^{z_j}}\right) = -S(z_i) \times S(z_j)$$

$$\frac{\partial \hat{y}}{\partial z^3} = (-y_1 y_2) + (-y_1 y_2)$$

$$= -2(y_1 y_2)$$

$$\frac{\partial z^3}{\partial w^3} = a^2$$

$$\frac{\partial L}{\partial w^3} = \left[-2(y_1 - \hat{y}_1) - 2(y_2 - \hat{y}_2)\right]\left[-2 y_1 y_2\right] a^2$$

$$= 4 y_1 y_2 \left[(y_1 - \hat{y}_1) + (y_2 - \hat{y}_2)\right] a^2$$

$$\frac{\partial L}{\partial w^2} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial z^3} \frac{\partial z^3}{\partial a^2} \frac{\partial a^2}{\partial z^2} \frac{\partial z^2}{\partial w^2}$$

$$\frac{\partial L}{\partial \hat{y}} = -2(y_1 - \hat{y}_1) - 2(y_2 - \hat{y}_2)$$

$$\frac{\partial \hat{y}}{\partial z^3} = (-y_1 y_2) + (-y_1 y_2) = -2(y_1 y_2)$$

$$\frac{\partial z^3}{\partial a^2} = \frac{\partial}{\partial a^2}(W^3 a^2) = W^3$$

$$\frac{\partial a^2}{\partial z^2} = \frac{\partial}{\partial z^2}(\sigma(z^2)) = a^2(1-a^2)$$

$$\frac{\partial z^2}{\partial w^2} = a^1$$

$$\frac{\partial L}{\partial w^2} = [-2(y_1 - \hat{y}_1) - 2(y_2 - \hat{y}_2)][-2(y_1 y_2)][w^3]$$
$$[a^2(1-a^2)]\,[a^1]$$

$$= -4 y_1 y_2 [(y_1 - \hat{y}_1) + (y_2 - \hat{y}_2)][w^3][a^2(1-a^2)][a^1]$$

$$\frac{\partial L}{\partial w^1} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z^3} \cdot \frac{\partial z^3}{\partial a^2} \cdot \frac{\partial a^2}{\partial z^2} \cdot \frac{\partial z^2}{\partial a^1} \cdot \frac{\partial a^1}{\partial z^1} \cdot \frac{\partial z^1}{\partial w^1}$$

$$\frac{\partial L}{\partial \hat{y}} = -2(y_1 - \hat{y}_1)(-2(y_2 - \hat{y}_2))$$

$$\frac{\partial \hat{y}}{\partial z^3} = -2(y_1 y_2)$$

$$\frac{\partial z^3}{\partial a^2} = W^3$$

$$\frac{\partial a^2}{\partial z^2} = a^2(1-a^2)$$

$$\frac{\partial z^2}{\partial a^1} = \frac{\partial}{\partial a^1}(W^2 a^1) = W^2$$

$$\frac{\partial a^1}{\partial z^1} = \frac{\partial}{\partial z^1}(\sigma(z^1)) = a^1(1-a^1)$$

$$\frac{\partial z^1}{\partial w^1} = X$$

$$\frac{\partial L}{\partial w^1} = [-2(y_1 - \hat{y}_1) - 2(y_2 - \hat{y}_2)][-2(y_1 y_2)][W^3][a^2(1-a^2)]$$
$$[W^2][a^1(1-a^1)][X]$$

$$= 4y_1 y_2[(y_1 - \hat{y}_1) + (y_2 - \hat{y}_2)][W^3][a^2(1-a^2)][W^2]$$
$$[a^1(1-a^1)][X]$$