CMSC 471 HW-5

1)

a) Input: 4
output: 1

b) hidden: 4

c) $h_1 = (Wl_{1,1} \cdot x_1) + (Wl_{2,1} \cdot x_2) + (Wl_{3,1} \cdot x_3) + (Wl_{4,1} \cdot x_4)$

$h_2 = (Wl_{1,2} \cdot x_1) + (Wl_{2,2} \cdot x_2) + (Wl_{3,2} \cdot x_3) + (Wl_{4,2} \cdot x_4)$

$h_3 = (Wl_{1,3} \cdot x_1) + (Wl_{2,3} \cdot x_2) + (Wl_{3,3} \cdot x_3) + (Wl_{4,3} \cdot x_4)$

$h_4 = (Wl_{1,4} \cdot x_1) + (Wl_{2,4} \cdot x_2) + (Wl_{3,4} \cdot x_3) + (Wl_{4,4} \cdot x_4)$

d) $out = W2_1[(Wl_{1,1} \cdot x_1) + (Wl_{2,1} \cdot x_2) + (Wl_{3,1} \cdot x_3) + (Wl_{4,1} \cdot x_4)] + W2_2[(Wl_{1,2} \cdot x_1) + (Wl_{2,2} \cdot x_2) + (Wl_{3,2} \cdot x_3) + (Wl_{4,2} \cdot x_4)] + W2_3[(Wl_{1,3} \cdot x_1) + (Wl_{2,3} \cdot x_2) + (Wl_{3,3} \cdot x_3) + (Wl_{4,3} \cdot x_4)] + W2_4[(Wl_{1,4} \cdot x_1) + (Wl_{2,4} \cdot x_2) + (Wl_{3,4} \cdot x_3) + (Wl_{4,4} \cdot x_4)]$

e) $\frac{\partial loss}{\partial W2_1} = \frac{\partial(y - [(W2_1 \cdot h_1) + (W2_2 \cdot h_2) + (W2_3 \cdot h_3) + (W2_4 \cdot h_4)])}{\partial(W2_1)}$

$= \frac{\partial y}{\partial W2_1} - \frac{\partial W2_2 \cdot h_1}{\partial W2_1} - \frac{\partial W2_2 \cdot h_2}{\partial W2_1} - \frac{\partial W2_3 \cdot h_3}{\partial W2_1} - \frac{\partial W2_4 \cdot h_4}{\partial W2_1}$

$= -h_1 = -(Wl_{1,1} \cdot x_1) - (Wl_{2,1} \cdot x_2) - (Wl_{3,1} \cdot x_3) - (Wl_{4,1} \cdot x_4)$

f) $h_1 = \text{Sigmoid}(W_{11}x_1 + W_{21}x_2 + W_{31}x_3 + W_{41}x_4)$

$h_2 = \text{sigmoid}(W_{12}x_1 + W_{22}x_2 + W_{32}x_3 + W_{42}x_4)$

$h_3 = \text{sigmoid}(W_{13}x_1 + W_{23}x_2 + W_{33}x_3 + W_{43}x_4)$

$h_4 = \text{sigmoid}(W_{14}x_1 + W_{24}x_2 + W_{34}x_3 + W_{44}x_4)$

$\text{Output} = W2_1h_1 + W2_2h_2 + W2_3h_3 + W2_4h_4$

g) activation function allows for nonlinearity, which is better for developing complex models

2) [11, 25, 34, 17, 18, 65, 23, 41, 73, 19, 7, 82], centroids: 15 and 55 (c_1, c_2)

a) C_1: [11, 25, 34, 17, 18, 23, 19, 7]
C_2: [65, 41, 73, 82]

b) $c_1 = \frac{\text{sum}(C_1)}{\text{len}(C_1)} = \frac{154}{8} = 19.25 = c_1$ (centroid 1)

$c_2 = \frac{\text{sum}(C_2)}{\text{len}(C_2)} = \frac{261}{4} = 65.25$

c) $C_1 = [11, 25, 34, 17, 18, 23, 41, 19, 7]$

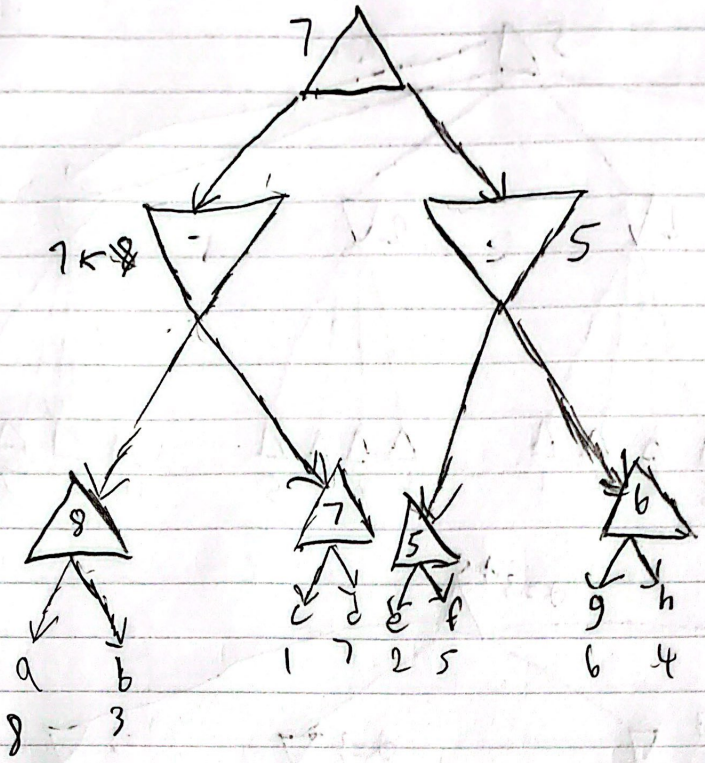
$C_2 = [65, 73, 82]$

d) - Only works for numeric data
- sensitive to input data
- must manually choose K

e) by having the number of possible values of a certain label or number of labels determine the number of clusters to create

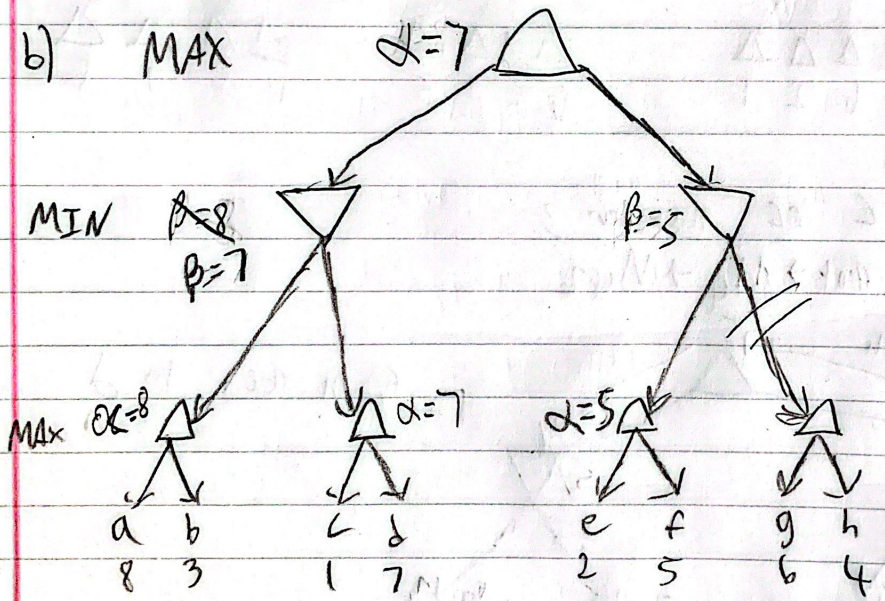
3)

normal



a) 7 and 5

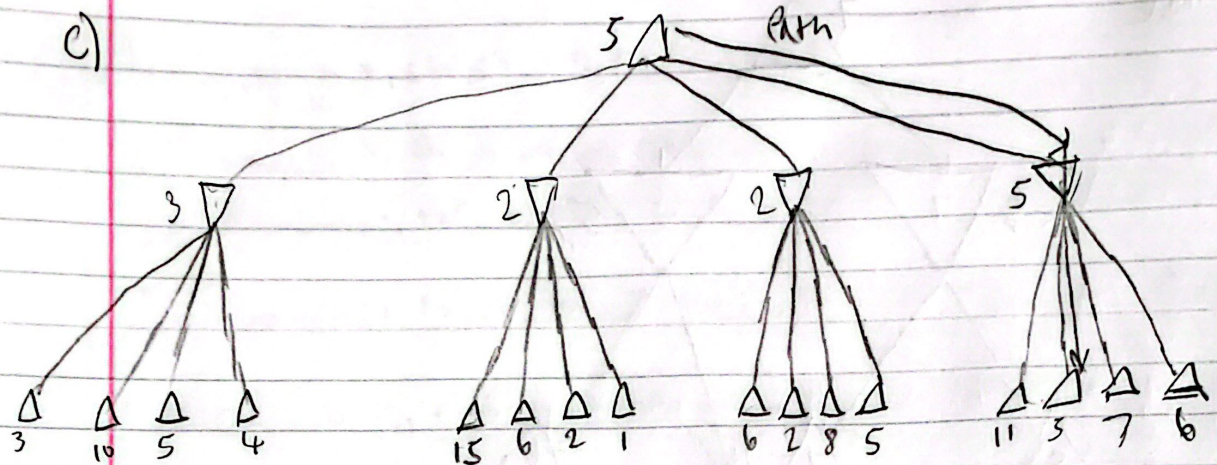
b) MAX



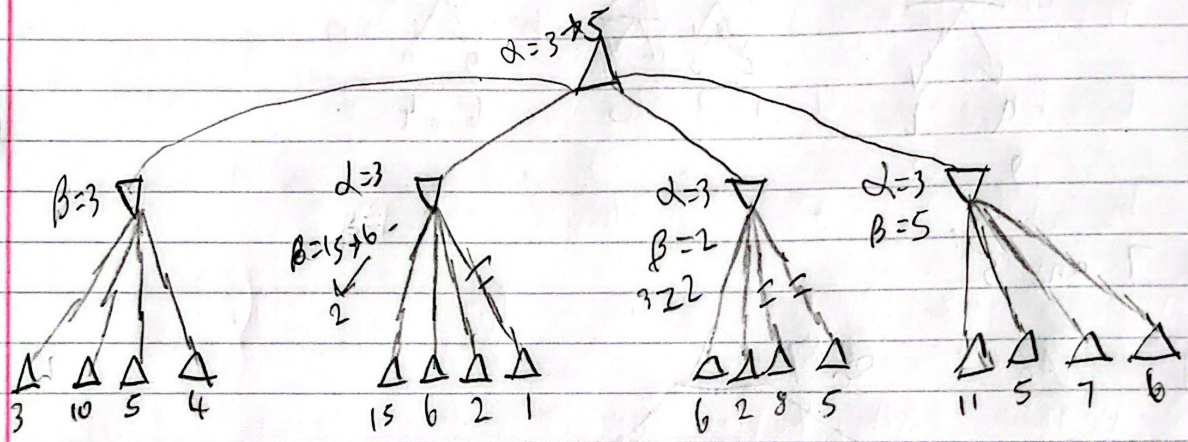
α ≥ β
7 ≥ 5

prunes values are 6 and 4

c)



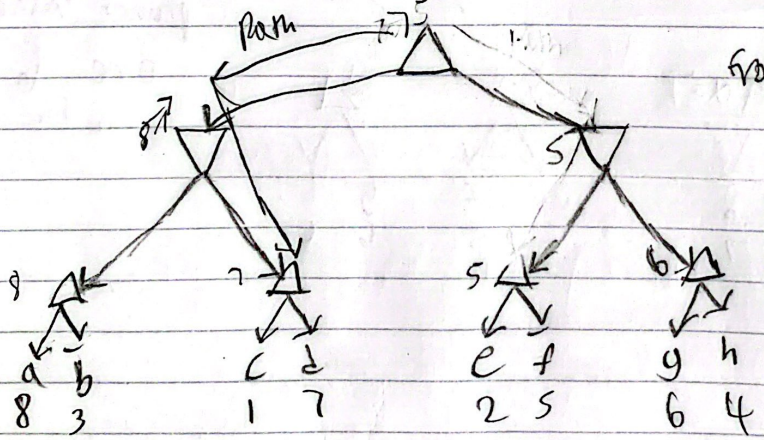
d)



e) Path drawn in ~~Part~~ C of number 3

It goes from root → M_4 → M_{42}

f)



from root to d