

a) Center for each cluster: $A1, A4, A7$, $K=3$
 c_1 c_2 c_3

1st Iteration:

$A1: \begin{aligned} d(A1, A1) &= 0 \\ d(A1, A4) &= \sqrt{3} \Rightarrow C_1 \\ d(A1, A7) &= \sqrt{65} \end{aligned}$

$A2: \begin{aligned} d(A2, A1) &= \sqrt{25} \\ d(A2, A4) &= \sqrt{18} \Rightarrow C_3 \\ d(A2, A7) &= \sqrt{10} \end{aligned}$

$A3: \begin{aligned} d(A3, A1) &= \sqrt{72} \\ d(A3, A4) &= \sqrt{25} \Rightarrow C_2 \\ d(A3, A7) &= \sqrt{53} \end{aligned}$

$A4: \begin{aligned} d(A4, A1) &= \sqrt{3} \\ d(A4, A4) &= 0 \Rightarrow C_2 \\ d(A4, A7) &= \sqrt{52} \end{aligned}$

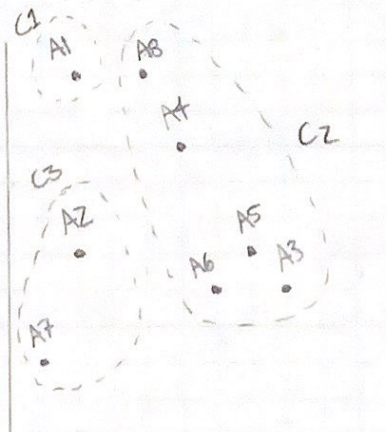
$A5: \begin{aligned} d(A5, A1) &= \sqrt{50} \\ d(A5, A4) &= \sqrt{13} \Rightarrow C_2 \\ d(A5, A7) &= \sqrt{45} \end{aligned}$

$A6: \begin{aligned} d(A6, A1) &= \sqrt{52} \\ d(A6, A4) &= \sqrt{7} \Rightarrow C_2 \\ d(A6, A7) &= \sqrt{29} \end{aligned}$

$A7: \begin{aligned} d(A7, A1) &= \sqrt{65} \\ d(A7, A4) &= \sqrt{52} \Rightarrow C_3 \\ d(A7, A7) &= 0 \end{aligned}$

$A8: \begin{aligned} d(A8, A1) &= \sqrt{5} \\ d(A8, A4) &= \sqrt{2} \Rightarrow C_2 \\ d(A8, A7) &= \sqrt{50} \end{aligned}$

$C_1: \{A1\}$; $C_2: \{A3, A4, A5, A6, A8\}$; $C_3: \{A2, A7\}$



b) $C1 = (2, 10)$

$$C2 = \left(\frac{8+5+7+6+4}{5}, \frac{4+8+5+4+9}{5} \right) = (6, 6) \Rightarrow$$

$$C3 = \left(\frac{3+1}{2}, \frac{5+7}{2} \right) = \left(\frac{3}{2}, \frac{7}{2} \right)$$

$C1 = (2, 10)$

$C2 = (6, 6)$

$C3 = \left(\frac{3}{2}, \frac{7}{2} \right)$

c) A1: $\partial(A1, C1) = 0$
 $\partial(A1, C2) = 4\sqrt{2} \Rightarrow C1$
 $\partial(A1, C3) = 6.52$

A2: $\partial(A2, C1) = \sqrt{26}$
 $\partial(A2, C2) = \sqrt{17} \Rightarrow C3$
 $\partial(A2, C3) = 1.58$

A3: $\partial(A3, C1) = \sqrt{72}$
 $\partial(A3, C2) = 2.82 \Rightarrow C2$
 $\partial(A3, C3) = 6.52$

A4: $\partial(A4, C1) = \sqrt{13}$
 $\partial(A4, C2) = \sqrt{5} \Rightarrow C2$
 $\partial(A4, C3) = 5.7$

A5: $\partial(A5, C1) = \sqrt{50}$
 $\partial(A5, C2) = \sqrt{2} \Rightarrow C2$
 $\partial(A5, C3) = 5.7$

A6: $\partial(A6, C1) = \sqrt{52}$
 $\partial(A6, C2) = 2 \Rightarrow C2$
 $\partial(A6, C3) = 4.8$

A7: $\partial(A7, C1) = \sqrt{16}$
 $\partial(A7, C2) = \sqrt{41} \Rightarrow C3$
 $\partial(A7, C3) = 1.158$

A8: $\partial(A8, C1) = \sqrt{5}$
 $\partial(A8, C2) = \sqrt{13} \Rightarrow C1$
 $\partial(A8, C3) = 6.04$

$$C1 = \left(\frac{2+4}{2}, \frac{9+0}{2} \right) = (3, 4.5)$$

$$C2 = \left(\frac{8+5+7+6}{4}, \frac{4+3+5+4}{4} \right) = (6.5, 5.25)$$

$$C3 = (1.5, 3.5)$$

A1: $\partial(A1, C1) = 1.12$
 $\partial(A1, C2) = 6.54 \Rightarrow C1$
 $\partial(A1, C3) = 6.52$

A2: $\partial(A2, C1) = 4.61$
 $\partial(A2, C2) = 4.51 \Rightarrow C3$
 $\partial(A2, C3) = 1.58$

A3: $\partial(A3, C1) = 7.43$
 $\partial(A3, C2) = 1.95 \Rightarrow C2$
 $\partial(A3, C3) = 6.52$

A4: $\partial(A4, C1) = 2.5$
 $\partial(A4, C2) = 3.13 \Rightarrow C1$
 $\partial(A4, C3) = 5.7$

A5: $\partial(A5, C1) = 6.02$
 $\partial(A5, C2) = 0.56 \Rightarrow C2$
 $\partial(A5, C3) = 5.7$

A6: $\partial(A6, C1) = 6.26$
 $\partial(A6, C2) = 1.35 \Rightarrow C2$
 $\partial(A6, C3) = 4.5$

A7: $\partial(A7, C1) = 7.76$
 $\partial(A7, C2) = 6.39 \Rightarrow C3$
 $\partial(A7, C3) = 1.158$

A8: $\partial(A8, C1) = 1.12$
 $\partial(A8, C2) = 4.51 \Rightarrow C1$
 $\partial(A8, C3) = 6.04$

$$C1 = \left(\frac{2+4+7}{3}, \frac{10+8+9}{3} \right) = (3.67, 9)$$

$$C2 = \left(\frac{8+7+6}{3}, \frac{4+8+4}{3} \right) = (7, 4.33)$$

$$C3 = (1.5, 3.5)$$

$$A1: d(A1, C1) = 1.95 \\ d(A1, C2) = 7.6 \Rightarrow C1 \\ d(A1, C3) = 6.52$$

$$A3: d(A3, C1) = 6.61 \\ d(A3, C2) = 1.65 \Rightarrow C2 \\ d(A3, C3) = 6.52$$

$$A5: d(A5, C1) = 5.2 \\ d(A5, C2) = 0.67 \Rightarrow C2 \\ d(A5, C3) = 5.7$$

$$A7: d(A7, C1) = 7.8 \\ d(A7, C2) = 6.4 \Rightarrow C3 \\ d(A7, C3) = 1.58$$

$$A2: d(A2, C1) = 4.33 \\ d(A2, C2) = 5.04 \Rightarrow C3 \\ d(A2, C3) = 1.58$$

$$A4: d(A4, C1) = 1.66 \\ d(A4, C2) = 4.2 \Rightarrow C1 \\ d(A4, C3) = 5.7$$

$$A6: d(A6, C1) = 5.52 \\ d(A6, C2) = 1.05 \Rightarrow C2 \\ d(A6, C3) = 4.5$$

$$A8: d(A8, C1) = 6.23 \\ d(A8, C2) = 5.55 \Rightarrow C1 \\ d(A8, C3) = 6.04$$

No change,

✓

2nd Iteration: $C1: \{A1, A8\}, C2: \{A3, A4, A5, A6\}, C3: \{A2, A7\}$

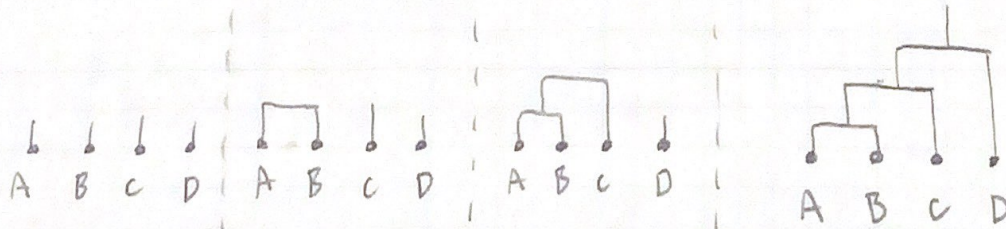
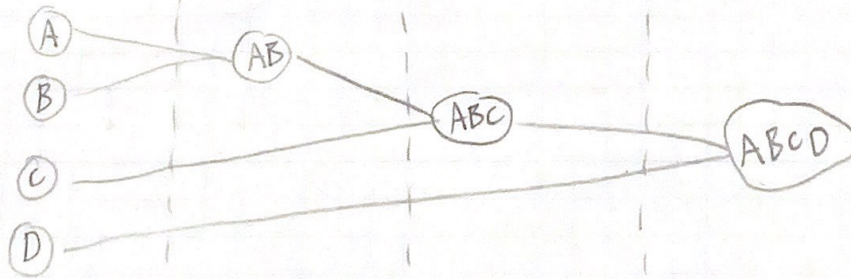
$$C1 = (3, 9.5), C2 = (6.5, 5.25), C3 = (1.5, 3.5)$$

3rd Iteration: $C1: \{A1, A4, A8\}, C2: \{A3, A5, A6\}, C3: \{A2, A7\}$

$$C1 = (3.67, 9), C2 = (7, 4.33), C3 = (1.5, 3.5)$$

- Needs 2 more Iteration to Converge.

(2) Step 0: Step 1: Step 2 Step 3



③

Class	Feature 1 (F ₁)	Feature 2 (F ₂)	Feature 3 (F ₃)
A	N	N	Y
A	N	Y	Y
A	N	Y	Y
A	N	Y	Y
A	N	Y	Y
A	Y	N	N
A	Y	N	Y
A	Y	Y	Y
A	Y	Y	Y
B	N	N	N
B	N	N	N
B	N	N	N
B	N	Y	N
B	Y	N	N
B	Y	N	Y
B	Y	Y	N
B	Y	Y	N
B	Y	N	N
B	Y	N	N

a) Gain F₁:

F ₁	A	B	I(A,B)
Y	4	6	0.971
N	5	4	0.991

$$E(F_1) = \frac{10}{19} I(4,6) + \frac{9}{19} I(5,4) = 0.9805$$

$$I(10,9) = 0.998$$

$$\text{Gain}(F_1) = I(10,9) - E(F_1) = 0.0175$$

Gain F₂:

F ₂	A	B	I(A,B)
Y	6	2	0.911
N	3	8	0.845

$$E(F_2) = \frac{8}{19} I(6,2) + \frac{11}{19} I(3,8) = 0.831$$

$$\text{Gain}(F_2) = I(10,9) - E(F_2) = 0.167$$

Gain F₃:

F ₃	A	B	I(A,B)
Y	8	2	0.722
N	1	8	0.503

$$E(F_3) = \frac{10}{19} I(8,2) + \frac{9}{19} I(1,8) = 0.618$$

$$\text{Gain}(F_3) = I(10,9) - E(F_3) = 0.380$$

// Now find information gain for F₁, F₂ given state of F₃

F₃ = Y:

F ₁	A	B	I(A,B)
Y	3	1	0.9113
N	5	1	0.6500

$$E(F_1) = \frac{4}{10} I(3,1) + \frac{6}{10} I(5,1) = 0.71532$$

$$I(8,2) = 0.7219$$

$$\text{Gain}(F_1) = 0.00661$$

F2	A	B	I(A,B)
Y	6	1	0.592
N	2	1	0.918

$$E(F2) = \frac{7}{10} I(6,1) + \frac{3}{10} I(2,1) = 0.6899$$

$$\text{Gain}(F2) = 0.03212$$

F3 = N:

F1	A	B	I(A,B)
Y	1	3	0.9544
N	0	5	0

$$E(F1) = \frac{4}{9} I(1,3) = 0.4242$$

$$I(1,8) = 0.50326$$

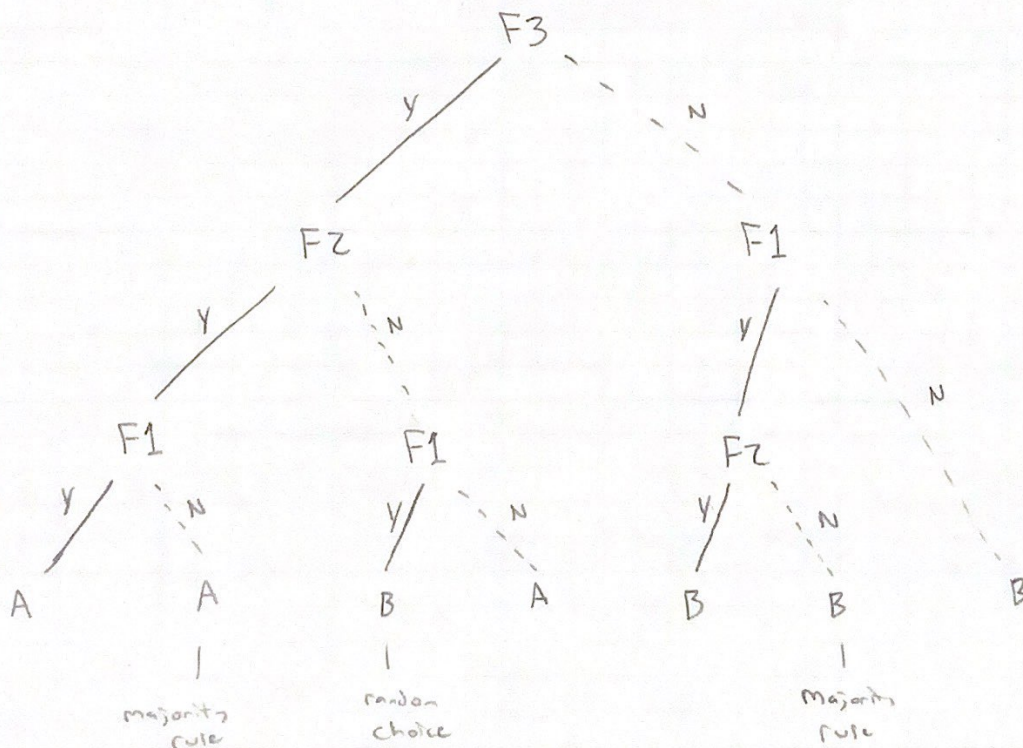
$$\text{Gain}(F1) = 0.079065$$

F2	A	B	I(A,B)
Y	0	1	0
N	1	7	0.54356

$$E(F2) = \frac{8}{9} I(1,7) = 0.483163$$

$$\text{Gain}(F2) = 0.0200$$

Decision Tree:



$$b) X = (F1=Y, F2=N, F3=N)$$

$$P(F1=Y | \text{class}=A) = 4/9$$

$$P(F1=N | \text{class}=A) = 5/9, P(\text{class}=A) = 9/19$$

$$P(F1=Y | \text{class}=B) = 6/10$$

$$P(F1=N | \text{class}=B) = 4/10, P(\text{class}=B) = 10/19$$

$$P(F2=N | \text{class}=A) = 3/9$$

$$P(F2=Y | \text{class}=A) = 6/9$$

$$P(F2=N | \text{class}=B) = 8/10$$

$$P(F2=Y | \text{class}=B) = 2/10$$

$$P(F3=N | \text{class}=A) = 1/9$$

$$P(F3=Y | \text{class}=A) = 8/9$$

$$P(F3=N | \text{class}=B) = 8/10$$

$$P(F3=Y | \text{class}=B) = 2/10$$

$$\text{Sample 1} = (F1=N, F2=N, F3=Y)$$

$$P(S1 | \text{class}=A) = (5/9)(3/9)(8/9) = \frac{40}{513}$$

$$P(S1 | \text{class}=B) = (4/10)(8/10)(2/10) = \frac{16}{475}$$

$$P(\text{class}=A | S1) = 0.698, P(\text{class}=B | S1) = 0.302$$

$$\text{Sample 2} = (F1=N, F2=Y, F3=Y)$$

$$P(S2 | \text{class}=A) = (5/9)(6/9)(8/9) = \frac{80}{513}$$

$$P(S2 | \text{class}=B) = (4/10)(2/10)(2/10) = \frac{4}{475}$$

$$P(\text{class}=A | S2) = 0.949, P(\text{class}=B | S2) = 0.051$$

$$\text{Sample 3} = (F1=N, F2=Y, F3=Y)$$

// Same as S2

$$\text{Sample 4} = (F1=N, F2=Y, F3=Y)$$

// Same as S2

$$\text{Sample 5} = (F1=N, F2=Y, F3=Y)$$

// Same as S2

$$X = (F1=Y, F2=N, F3=N)$$

$$P(X | \text{class}=A) = (4/9)(3/9)(1/9) = 0.0078$$

$$P(X | \text{class}=B) = (6/10)(8/10)(8/10) = 0.202$$

\Rightarrow X belongs to class B