Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |
Last Name: Leos
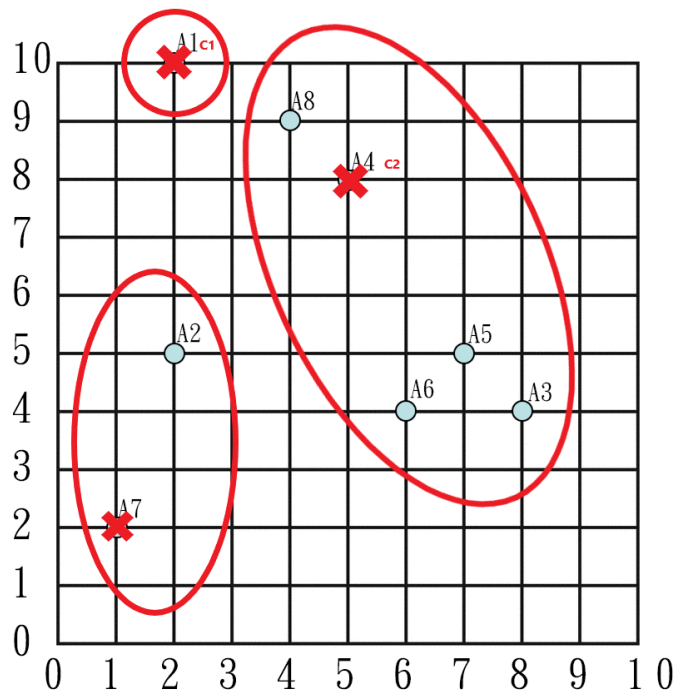First Name: Hugo

Assignment 5

1.

a.

Iteration 1								
Centroid(C1 = A1, C2 = A4, C3 = A7)								
Instance	A1	A2	A3	A4	A5	A6	A7	A8
C1 dist.	0.0000	5.0000	8.4853	3.6056	7.0711	7.2111	8.0623	2.2361
C2 dist.	3.6056	4.2426	5.0000	0.0000	3.6056	4.1231	7.2111	1.4142
C3 dist.	8.0623	3.1623	7.2801	7.2111	6.7082	5.3852	0.0000	7.6158
Cluster Assigned	C1	C3	C2	C2	C2	C2	C3	C2

	x	y
C1	2	10
C2	6	6
C3	1.5	3.5

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |
Last Name: Leos
First Name: Hugo

Iteration 2								
Centroid(C1, C2, C3)								
Instance	A1	A2	A3	A4	A5	A6	A7	A8
C1 dist.	0.0000	5.0000	8.4853	3.6056	7.0711	7.2111	8.0623	2.2361
C2 dist.	5.6569	4.1231	2.8284	2.2361	1.4142	2.0000	6.4031	3.6056
C3 dist.	6.5192	1.5811	6.5192	5.7009	5.7009	4.5277	1.5811	6.0415
Cluster Assigned	C1	C3	C2	C2	C2	C2	C3	C1
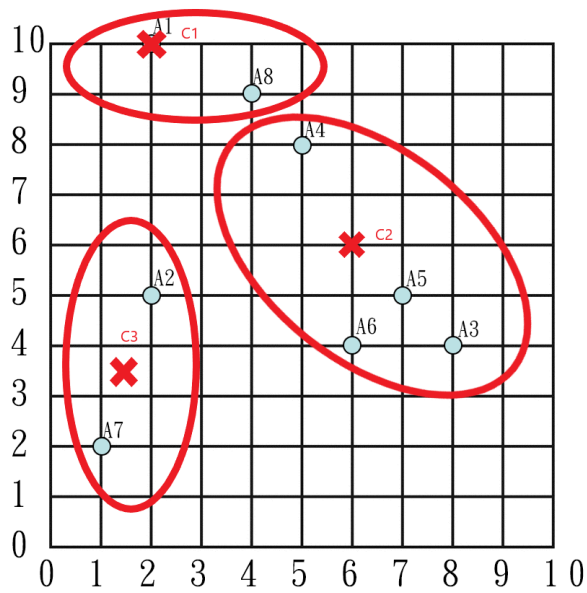
	x	y
C1	3	9.5
C2	6.5	5.25
C3	1.5	3.5

Iteration 3								
Centroid(C1, C2, C3)								
Instance	A1	A2	A3	A4	A5	A6	A7	A8
C1 dist.	1.1180	4.6098	7.4330	2.5000	6.0208	6.2650	7.7621	1.1180
C2 dist.	6.5431	4.5069	1.9526	3.1325	0.5590	1.3463	6.3885	4.5069
C3 dist.	6.5192	1.5811	6.5192	5.7009	5.7009	4.5277	1.5811	6.0415
Cluster Assigned	C1	C3	C2	C1	C2	C2	C3	C1

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |
Last Name: Leos
First Name: Hugo
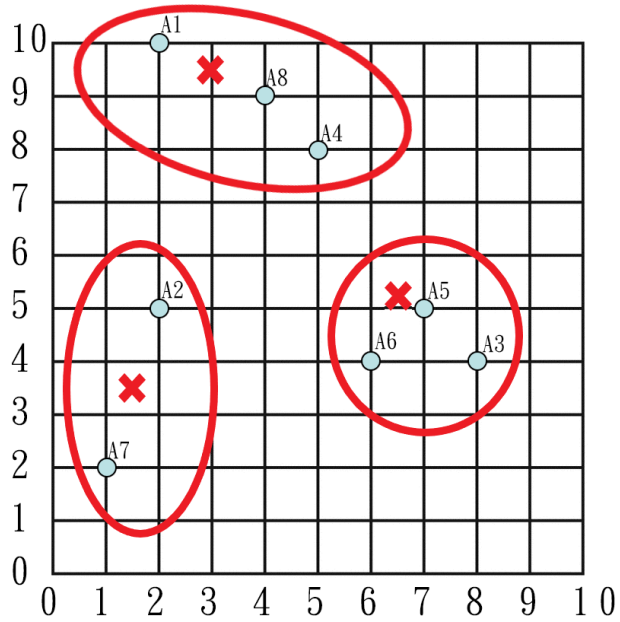
	x	y
C1	3.6667	9
C2	7	4.3333
C3	1.5	3.5

Iteration 4								
Centroid(C1, C2, C3)								
Instance	A1	A2	A3	A4	A5	A6	A7	A8
C1 dist.	1.9437	4.3333	6.6165	1.6667	5.2068	5.5176	7.4907	0.3333
C2 dist.	7.5572	5.0442	1.0541	4.1767	0.6667	1.0541	6.4377	5.5478
C3 dist.	6.5192	1.5811	6.5192	5.7009	5.7009	4.5277	1.5811	6.0415
Cluster Assigned	C1	C3	C2	C1	C2	C2	C3	C1

No change = convergence.

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |
Last Name: Leos
First Name: Hugo

b.

$$SSE = \left(dist^2(C1, A1) + dist^2(C1, A4) + dist^2(C1, A8)\right)$$
$$+ \left(dist^2(C2, A3) + dist^2(C2, A5) + dist^2(C2, A6)\right)$$
$$+ \left(dist^2(C3, A2) + dist^2(C3, A7)\right)$$

$$SSE = ((1.9437)^2 + (1.6667)^2 + (0.3333)^2) + ((1.0541)^2 + (0.6667)^2 + (1.0541)^2)$$
$$+ ((1.5811)^2 + (1.5811)^2)$$

$$SSE = 14.3334$$

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |
Last Name: Leos
First Name: Hugo

2. Link

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |
Last Name: Leos
First Name: Hugo

   3.

       a.

			User-based			
	r_i • r_k	||r_i|| ||r_k||	sim(u_k, u_i)		prediction Kiss	prediction GnR
f(Fred, John)	18	21.12463017	0.852085923		2.3301714	1.552984631
f(Lillian, John)	20	20.10596926	0.994729463			
f(Cathy, John)	20	21	0.952380952			

For Kiss, there is a predicted rating of 2.33 > 2.0 therefore it is recommended.

For GnR, the predicted rating is 1.55 < 2.0 therefore it is not recommended.

       b.

			Item-based		
	r_i • r_k	||r_i|| ||r_k||	sim(u_k, u_i)		prediction Kiss
f(Bon Jovi, Kiss)	13	13.09580085	0.992684613		2.4286809
f(Metallica, Kiss)	10.5	13.66793693	0.76822128		
f(Scorpions, Kiss)	10.5	11.20546742	0.937042571		
f(AC/DC, Kiss)	13.5	16.41645516	0.82234562		

For Kiss, the predicted rating is 2.43 > 2.0 therefore it is recommended.

			Item-based		
	r_i • r_k	||r_i|| ||r_k||	sim(u_k, u_i)		prediction GnR
f(Bon Jovi, GnR)	10	14	0.714285714		1.586354
f(Metallica, GnR)	14.5	14.6116392	0.992359571		
f(Scorpions, GnR)	10.5	11.97914855	0.876523065		
f(AC/DC, GnR)	17	17.54992877	0.9686649		

For GnR, the predicted rating is 1.59 < 2.0 therefore it is not recommended.

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |
Last Name: Leos
First Name: Hugo

4. 4

    a.

1-Itemsets			
Rank	Itemset	Support Count	Support
1	d	9	0.9
2	b	7	0.7
3	e	6	0.6
4	a	5	0.5
5	c	5	0.5

2-Itemsets			
Rank	Itemset	Support Count	Support
1	{b, d}	6	0.6
2	{d, e}	6	0.6
3	{a, d}	4	0.4
4	{a, e}	4	0.4
5	{b, e}	4	0.4
6	{c, d}	4	0.4
7	{a, b}	3	0.3
8	{b, c}	3	0.3
9	{a, c}	2	0.2
10	{c, e}	2	0.2

3-Itemsets			
Rank	Itemset	Support Count	Support
1	{a, d, e}	4	0.4
2	{b, d, e}	4	0.4
3	{a, b, d}	2	0.2
4	{a, b, e}	2	0.2
5	{b, c, d}	2	0.2
6	{a, b, c}	1	0.1

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |
Last Name: Leos
First Name: Hugo

b.

Rule Generation				
Rank	{a, d, e}	Confidence	{b, d, e}	Confidence
1	AD -> E	1	BE -> D	1
2	AE -> D	1	BD -> E	0.666666667
3	A -> DE	0.8	DE -> B	0.666666667
4	DE -> A	0.666666667	E -> BD	0.666666667
5	E -> AD	0.666666667	B -> DE	0.571428571
6	D -> AE	0.444444444	D -> BE	0.444444444

c.

3-Itemsets		
Itemset	Support Count	Support
ABD	2	0.2
ABE	2	0.2
BCD	2	0.2
BCE	1	0.1
BDE	4	0.4

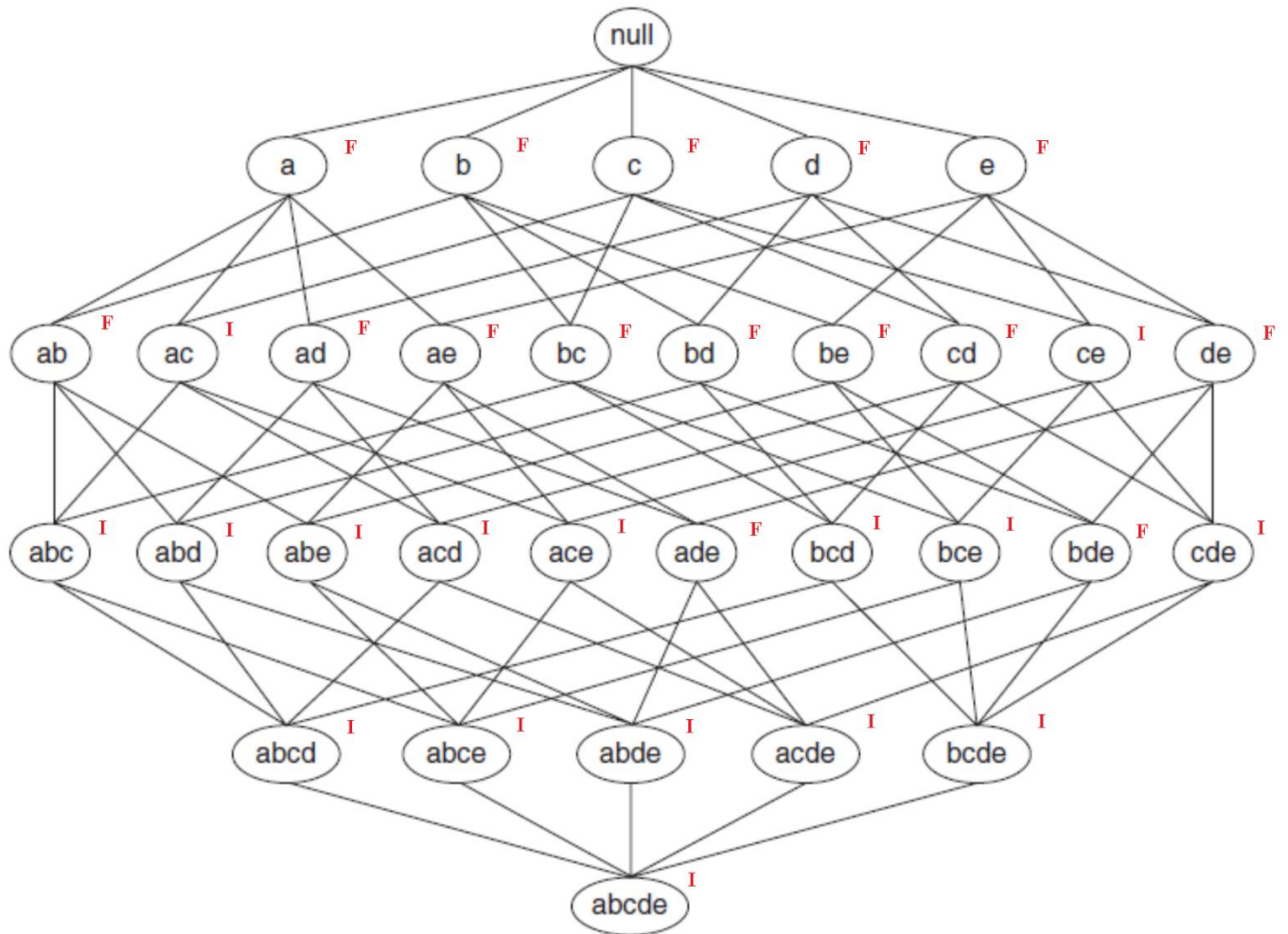| 3-Itemsets		
Itemset	Support Count	Support
BDE	4	0.4

All except BDE get pruned.

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |
Last Name: Leos
First Name: Hugo

d.

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |
Last Name: Leos
First Name: Hugo

5. Link