Name : Pawan Suresh Bathe Roll No.:CS18M519

CS6745: Mining Massive DataSets

Tutorial 9

November 13, 2019

- Write your name and roll number in the space provided
- Be neat, and use the space judiciously.
- **Rough sheets won't be evaluated.**

1. Suppose there are 100 items, numbered 1 to 100, and also 100 baskets, also numbered 1 to 100. Item i is in basket b if and only if i divides b with no remainder. Thus, item 1 is in all the baskets, item 2 is in all fifty of the even-numbered baskets, and so on. Basket 12 consists of items 1, 2, 3, 4, 6, 12, since these are all the integers that divide 12. Answer the following questions,

   (a) (2 marks) If the support threshold is 5, which items are frequent?

   Ans: Items {1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20} appears in at-least 5 buckets.
   So, If the support threshold is 5, frequent items are:
   {1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20}.

   (b) (3 marks) If the support threshold is 5, which pairs of items are frequent?

   Ans: For the set {1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20} in Q1.

   Pairs which will appear in at-least 5 buckets and hence frequents items pairs if support threshold is 5 are:

   {(1,2),(1,3),(1,4),(1,5),(1,6),(1,7),(1,8),(1,9),(1,10),(1,11),(1,12),(1,13),(1,14),(1,15),(1,16),
   (1,17),(1,18), (1,19),(1,20),(2,4),(2,6),(2,8),(2,10),(2,12),(2,14),(2,16),(2,18),(2,20),(3,6),
   (3,9),(3,12),(3,15),(3,18),(4,8), (4,12),(4,16),(4,20),(5,10),(5,15),(5,20),(6,12),(6,18),(7,14),
   (8,16),(9,18),(10,20)}

2. (5 marks) Apply the A-Priori Algorithm with support threshold 5 to the data present in Q1.

   Ans:
   Possible items set of size 1:
   1,2,3,...,100
   Actual frequent items set of size 1 with support threshold of 5 are:
   1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20

   Possible items set of size 2:
   Any two elements pairs from 1 element frequent size elements.

Actual frequent items set of size 2 with support threshold of 5 are:
{(1,2),(1,3),(1,4),(1,5),(1,6),(1,7),(1,8),(1,9),(1,10),(1,11),(1,12),(1,13),(1,14),(1,15), (1,16), (1,17), (1,18), (1,19),(1,20),(2,4),(2,6),(2,8),(2,10),(2,12),(2,14),(2,16),(2,18), (2,20),(3,6), (3,9),(3,12), (3,15),(3,18),(4,8), (4,12),(4,16),(4,20),(5,10),(5,15),(5,20), (6,12),(6,18),(7,14), (8,16),(9,18), (10,20)}

Possible items set of size 3:
Any three elements sets formed from elements of sets in two items set.
Actual frequent items set of size 3 with support threshold of 5 are:
{(1, 2, 3), (1, 2, 4), (1, 2, 5), (1, 2, 6), (1, 2, 7), (1, 2, 8), (1, 2, 9), (1, 2, 10), (1, 2, 12), (1, 2, 14), (1, 2, 16), (1, 2, 18), (1, 2, 20), (1, 3, 4), (1, 3, 5), (1, 3, 6), (1, 3, 9), (1, 3, 12), (1, 3, 15), (1, 3, 18), (1, 4, 5), (1, 4, 6), (1, 4, 8), (1, 4, 10), (1, 4, 12), (1, 4, 16), (1, 4, 20), (1, 5, 10), (1, 5, 15), (1, 5, 20), (1, 6, 9), (1, 6, 12), (1, 6, 18), (1, 7, 14), (1, 8, 16), (1, 9, 18), (1, 10, 20), (2, 3, 4), (2, 3, 6), (2, 3, 9), (2, 3, 12), (2, 3, 18), (2, 4, 5), (2, 4, 6), (2, 4, 8), (2, 4, 10), (2, 4, 12), (2, 4, 16), (2, 4, 20), (2, 5, 10), (2, 5, 20), (2, 6, 9), (2, 6, 12), (2, 6, 18), (2, 7, 14), (2, 8, 16), (2, 9, 18), (2, 10, 20), (3, 4, 12), (3, 5, 15), (4, 5, 10), (4, 5, 20), (4, 6, 12), (5, 10, 20), (6, 9, 18)}

Possible items set of size 4:
Any 4 elements set formed from elements of 3 elements sets.
Actual frequent items set of size 4 with support threshold of 5 are:

{(1, 2, 3, 4), (1, 2, 3, 6), (1, 2, 3, 9), (1, 2, 3, 12), (1, 2, 3, 18), (1, 2, 4, 5), (1, 2, 4, 6), (1, 2, 4, 8), (1, 2, 4, 10), (1, 2, 4, 12), (1, 2, 4, 16), (1, 2, 4, 20), (1, 2, 5, 10), (1, 2, 5, 20), (1, 2, 6, 9), (1, 2, 6, 12), (1, 2, 6, 18), (1, 2, 7, 14), (1, 2, 8, 16), (1, 2, 9, 18), (1, 2, 10, 20), (1, 3, 4, 12), (1, 3, 5, 15), (1, 4, 5, 10), (1, 4, 5, 20), (1, 4, 6, 12), (1, 5, 10, 20), (1, 6, 9, 18) (2, 3, 4, 12), (2, 4, 5, 10), (2, 4, 5, 20), (2, 4, 6, 12), (2, 5, 10, 20), (2, 6, 9, 18) (4, 5, 10, 20)}

Possible items set of size 5:
Any 5 elements set formed from elements of 4 elements sets.
Actual frequent items set of size 5 with support threshold of 5 are:
{(1, 2, 3, 4, 12), (1, 2, 4, 5, 10), (1, 2, 4, 5, 20), (1, 2, 4, 6, 12), (1, 2, 5, 10, 20), (1, 2, 6, 9, 18), (1, 4, 5, 10, 20), (2, 4, 5, 10, 20)}

Possible items set of size 6:
Any 6 elements set formed from elements of 5 elements sets.
Actual frequent items set of size 6 with support threshold of 5 are:
{1, 2, 4, 5, 10, 20}

Since there is only one frequent items set, we can stop here.