Assignment 2

Sai Vishal Tyagadurgam

September 2020

1. Compute the support for items {e}, {b,d}, and {b,d,e} by treating each transaction ID as a market basket.

   Support for an item is a fraction of transactions that contain an itemset.
   support = number of transactions that contain an item / total number of transactions

   - Support for item {e} is $8/10 = 0.8$
   - Support for item {b,d} is $2/10 = 0.2$
   - Support for item {b,d,e} is $2/10 = 0.2$

2. Use the result in part (a) to compute the confidence for the association rules {b,d} $\rightarrow$ {e} and {e} $\rightarrow$ {b,d}. Is confidence a symmetric measure?

   In an expression like X $\rightarrow$ Y confidence measures how often items in Y appear in transactions that contains X.
   confidence = number of transactions that contain items in X, Y / number of transactions that contain items in X.

   - Confidence of the association rule {b,d} $\rightarrow$ {e} is $2/2 = 1$.
   - Confidence of the rule {e} $\rightarrow$ {b,d} is $2/8 = 0.25$

   From the above results we can conclude that confidence is asymmetric in nature.

3. Repeat part (a) by treating each customer ID as a market basket. Each item should be treated as a binary variable (1 if an item appears in at least one transaction bought by the customer, and 0 otherwise).

   By considering customer ID as a market basket the total number of transactions will be 5.

   - Support of item {e} is $4/5 = 0.8$
   - Support of item {b,d} is $5/5 = 1$

- Support of item {b,d,e} is 4/5 = 0.8

4. Use the results in part (c) to compute the confidence for the association rules {b,d} → {e} and {e} → {b,d}

    - Confidence of {b,d} → {e} is 4/5 = 0.8
    - Confidence of {e} → {b,d} is 4/4 = 1

5. Suppose s1 and c1 are the support and confidence values of an association rule r when treating each transaction ID as a market basket. Also, let s2 and c2 be the support and confidence values of r when treating each customer ID as a market basket. Discuss whether there are any relationships between s1 and s2 or c1 and c2.

Association Rule	Transaction ID as market basket	Customer ID as market basket
Support of item {e}	0.8	0.8
Support of item {b,d}	0.2	1
Support of item {b,d,e}	0.2	0.8
Confidence of rule {b,d} → {e}	1	0.8
Confidence of rule {e} → {b,d}	0.25	1

After observing the support and confidence for each transaction we can conclude there is no relationship between them while treating different ID as a market basket. These values changed with the change in ID.