Fundamentals of Data Science

Semester B 20-21

Tutorial 9

1. Consider the data set shown in the following table

Customer ID	Transaction ID	Items Bought
1	0001	{a,d,e}
1	0024	{a,b,c,e}
2	0012	{a,b,d,e}
2	0031	{a,c,d,e}
3	0015	{b,c,e}
3	0022	{b,d,e}
4	0029	{c,d}
4	0040	{a,b,c}
5	0033	{a,d,e}
5	0038	{a,b,e}

   a. Compute the support for itemsets {e}, {b,d} and {b,d,e} by treating each transaction ID as a market basket.
   b. Use the results in part a. to compute the confidence for the association rules {b,d}→{e} and {e}→{b,d}.
   c. Repeat part a. by treating each customer ID as a market basket.
   d. Use the results in part c. to compute the confidence for the association rules {b,d}→{e} and {e}→{b,d}.

2. Let c_1, c_2 and c_3 be the confidence values for the rules {p}→{q}, {p}→{q,r} and {p,r}→{q} respectively. If we assume that c_1, c_2 and c_3 have different values, what are the possible relationships that may exist among c_1, c_2 and c_3 ?