Assignment no.1
Intelligent data Analysis (IDA)
B Tech(h) Data Science

Date: 8th December,23

1. Explain association rule mining and list all the preliminary requirements for it.
2. For each of the following questions, provide an example of an association rule from the market basket domain that satisfies the following conditions. Also, describe whether such rules are subjectively interesting. (4M)

   a. A rule that has high support and high confidence.
   b. A rule that has reasonably high support but low confidence.
   c. A rule that has low support and low confidence.
   d. A rule that has low support and high confidence

3. Consider the data set shown

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

4. Discuss the Frequent item set generation and also explain antimonotone property
5. Elaborate maximal and closed frequent item sets. Also list the difference between them