

Fundamentals Data Mining
Data Science
Faculty of Computing
Sri Lanka Institute of Information Technology

Tutorial 01– Association Rule Mining

Question 1

Define the following terms for an association rule:

- (i) Support
- (ii) Confidence

Question 2

Explain why association rules cannot be directly used for prediction.

Question 3

Consider the following database instance D.

<i>transid</i>	<i>custid</i>	<i>date</i>	<i>item</i>	<i>qty</i>
111	201	5/1/99	pen	2
111	201	5/1/99	ink	1
111	201	5/1/91	milk	3
111	201	5/1/99	juice	6
112	105	6/3/99	pen	1
112	105	6/3/99	ink	1
112	105	6/3/99	milk	1
113	106	5/10/99	pen	1
113	106	5/10/99	milk	1
114	201	6/1/99	pen	2
114	201	6/1/99	ink	2
114	201	6/1/99	juice	4

- (a.) Calculate the frequent itemsets with minimum support 50% using the Apriori algorithm.
- (b.) Compute the confidence level for the following association rule:
 $\{juice\} \Rightarrow \{milk\}$

Question 4

Generation of the candidate itemsets and frequent itemset, where the minimum support count is 2.

<i>TID</i>	<i>List of item IDs</i>
T100	I1, I2, I5
T200	I2, I4
T300	I2, I3
T400	I1, I2, I4
T500	I1, I3
T600	I2, I3
T700	I1, I3
T800	I1, I2, I3, I5
T900	I1, I2, I3