Name: Student_ID:

Department of Computer Engineering

Faculty of Engineering

Kasetsart University

Exercise, Association Rule Mining

Exercise 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}*. 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}* → *{e}* and *{e}* → *{b, d}*.

Name: Student_ID:

Exercise 2. A database contains 6 transactions. Let minimum support = 50% and minimum confidence = 75%.

Trans_Id	Items
1	F,G,I,J
2	G,H,J
3	G,I,J
4	G,H,J
5	G,H,I,J
6	G,H,I

- Find all the frequent itemsets using Apriori. Please show every step Apriori (). For each frequent itemset, please indicate its support.
- Give 3 examples of strong association rules that are discovered from the database. For each association rule example, indicate its support & confidence values.

Exercise 3. Repeat the Exercise 2, using RapidMiner

- Find all the frequent itemsets using **FP-Growth** operator. For each frequent itemset, please indicate its support. => **Capture Screen**
- Give 3 examples of strong association rules that are discovered from the database. For each association rule example, indicate its support & confidence values. => **Capture Screen**
- Submit **Student_ID_hw#4.xml**