CSC 422/522: Automated Learning and Data Analysis
Homework 5: Due 4/23 @ 23:55

(No programming is required, do all work by hand, you can use calculator if needed)

Student Name:
Student ID:

Q1. Consider the market basket transactions given in Table 1, and answer all sub-questions (a)-(f) listed below. (1 point each, total 6 points)

TID	Items
1	Bread, Milk, Butter
2	Bread, Milk, Beer, Eggs
3	Milk, Diaper, Beer, Coke
4	Bread, Milk, Diaper, Beer
5	Bread, Milk, Diaper, Coke

Table 1. Market Basket Dataset

(a) What is the maximum number of association rules that can be extracted from the data given in Table 6 (including rules that have zero support)?

(b) What is the maximum size of frequent itemsets that can be extracted (assuming minsup > 0)?

(c) Compute the maximum number of size-3 itemsets that can be derived from this dataset

(d) What is the support of {bread}, {milk}, {bread, milk}?

(e) What is the confidence of the rule {bread} -> {milk} and {milk} -> {bread}

(f) Under what conditions the rules {a} -> {b} and {b} -> {a} have same confidence?

Q2. Apriori algorithm: Consider the dataset given in Table 2 and answer the sub-questions (a) and (b) using apriori algorithm. (total 10 points)

TID	Items
1	a, c, d
2	b, c, e
3	a, b, c, e
4	b, e

Table 2.

(**a**) Show (compute) each step of frequent itemset generation process using apriori algorithm, for support count of 2. (5 points).

(**b**) Show the lattice structure for the data given in Table 2, and mark the pruned branches if any. (5 points).

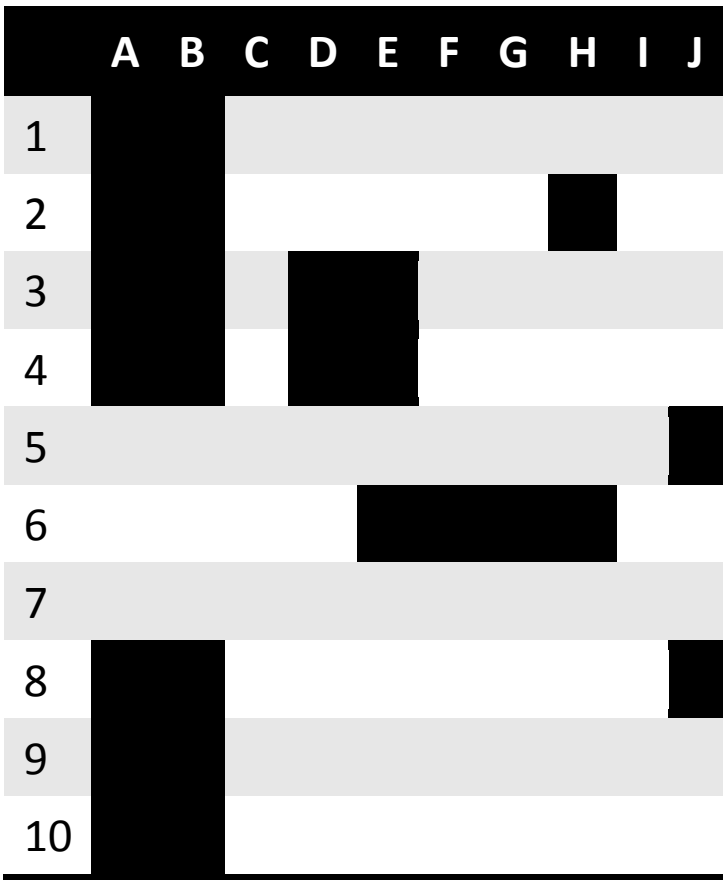
Q3. For the dataset given below, answer the following: (total 10 points)

(a) for support threshold of 6 (by support count), list frequent itemsets and maximal itemsets. (5 points)

(b) Repeat (a) for support thresholds of 5 and 4. (5 points)

	A	B	C	D	E	F	G	H	I	J
1										
2	■		■	■	■	■				■
3			■	■	■	■		■		
4			■	■	■	■				■
5					■	■				
6					■	■				
7					■	■				■
8		■	■							
9		■	■							■
10										

Q4. For the given dataset (10 items and 10 transactions), compute h-confidence for the itemset: {A,B,D} (4 points)



Practice (not graded)
Chapter 6.

Solve the following questions.
Q2
Q3
Q8
Q11