4

FIM - Quiz4

February 3, 2025

Name and surname: PABLO MOLLA
Instructions

- Each question has one and only one correct answer. Circle the corresponding letter.
- Grading: correct answer (1 pt), incorrect answer (-0.5 pt), no answer (0 pt).

1. Consider a dataset D_1 with the following transactions: $T_1 = ABC$, $T_2 = AB$, $T_3 = AC$, $T_4 = BC$, $T_5 = A$, $T_6 = B$, $T_7 = C$. Which of these itemsets are closed?
   A. A is closed, B and C are not.
   B. All of A, B, and C are closed.
   C. Only B is closed.
   D. None of A, B, or C are closed.

   t_1 A B C
   t_2 A B
   t_3 A A C
   B C
   t_4
   t_5 A
   t_6 B
   t_7 C

2. Consider again dataset D_1 with $\alpha = 2$. Which of the following itemsets are maximal?
   A. The itemset AB is maximal.
   B. All of A, B, and C are maximal.
   C. BC is not maximal.
   D. The itemset ABC is maximal.

3. Consider again dataset D_1 with $\alpha = 4$. Which of the following are the frequent closed itemsets?
   A. AC and BC.
   B. A, B, and C.
   C. $\emptyset$.
   D. $\emptyset$, A, B, and C.

4. What is the definition of a "closed itemset"?
   A. An itemset that occurs in at least 50
   B. An itemset that has the highest support in the dataset.
   C. An itemset for which there is no superset with the same support count.
   D. An itemset that is contained in every transaction.

5. Which of the following is true about maximal itemsets?
   A. A maximal itemset is always closed.
   B. A closed itemset is always maximal.
   C. A maximal itemset has no subset that is frequent.
   D. A maximal itemset has the highest support among all itemsets.

6. Which of the following is NOT an advantage of closed itemsets?
   A. They reduce redundancy in the output.
   B. They allow for lossless compression of frequent patterns.
   C. They always result in fewer itemsets than maximal itemsets.
   D. They maintain the same support counts as their supersets.

7. What is the "closure extension" rule in LCM used for?
   A. To generate candidate itemsets and compute their closure.
   B. To filter irrelevant transactions.
   C. To ensure that all itemsets have a support count greater than the minimum threshold.
   D. To extend closed itemsets by adding a new item and check if they remain closed.

8. What does the prefix of an itemset P represent in the prefix tree?
   A. The set of items in P that appear most frequently.
   B. The subset of items in P excluding the last item.
   C. The entire itemset P.
   D. The first item in P.

9. Which of the following best describes the time complexity of the LCM algorithm for enumerating closed itemsets?
   A. Linear time in the number of items in the dataset.
   B. Exponential time in the number of transactions.
   C. Polynomial time in the number of itemsets.
   D. Linear time in the number of frequent closed itemsets.

10. Why is the LCM algorithm particularly efficient for mining frequent closed itemsets?
   A. It avoids candidate generation by directly constructing maximal itemsets.
   B. It only considers itemsets with the highest support values.
   C. It uses a combination of prefix-preserving closure extension and efficient transaction pruning.
   D. It processes the dataset in a single pass without recursion.