TU Dortmund 21.4.2022
Artificial Intelligence / Data Science and Data Engineering **Machine Learning for Complex Data**
Prof. Dr. Erich Schubert
Prof. Dr. Emmanuel Müller **Summer 2022**
Daniel Wilmes

Assignment 2: Eclat and FP-Growth

Due: Thursday, 28.4.2022

Problem 2-1 Eclat algorithm **8**

Let the set of items be {A, B, C, D, E, F, G, H, I, K, L, M}.

The transactions *T* are given by the following table:

Transaction database *T*

T_ID	items bought	T_ID	items bought
1	B E G H	7	A B D G H
2	A B C E G H	8	A B D G
3	A B C E F H	9	B D F G
4	B C D E F G H L	10	C E F
5	A B E H K	11	A C E F H
6	B E F G H I K	12	A B E G

For the minimum support 30%, determine the frequent itemsets using the Eclat algorithm.

Problem 2-2 Association rules **5**

Determinte all association rules from the frequent itemset {B E G H} with a minimum confidence of 60% and a minimum support of 30%. Use the monotonicity discussed in the lecture to avoid redundant work by only pulling items to the right hand side that are alphabetically subsequent to any already on the right hand side.

Use the support information from the Apriori or Eclat assignments, do not count again.

Problem 2-3 FP-Growth algorithm **5**

Let the set of items be {A, B, C, D, E, F, G, H, I, K, L, M}.

The transactions *T* are given by the following table:

Transaction database *T*

T_ID	items bought	T_ID	items bought
1	B E G H	7	A B D G H
2	A B C E G H	8	A B D G
3	A B C E F H	9	B D F G
4	B C D E F G H L	10	C E F
5	A B E H K	11	A C E F H
6	B E F G H I K	12	A B E G

For the minimum support 30%, build the FP-Tree. If two items have the same frequency, sort them alphabetically. You do not need to draw a header table.

a) Perform the first pass of the FP-tree construction

b) Show the FP-Tree after only inserting the first half of the transactions.
Hint: verify that the total number of items agrees.

c) Show the FP-tree after also inserting the remaining transactions