TDT4300 - Data Mining

Assignment 2

Implementation of Apriori

by Marte Løge and Linn Vikre

1 Implementation of Frequent Itemset Methods

In this exercise we implemented 3 different methods for generating the frequent itemsets. We will in describe the implementation of each in this section.

1.1 Brute-Force

This method is quite simple; it genereates all possible combination and prune candidates that have support less than the threshold (minsupport), which is 0.4 in this implementation.

1.2 FKminus1F1

1.3 FKminus1FKminus1

2 Scaleability of Implemented Methods

2.1 Small Data Set

Method	Level 1	Level 2	Level 3	Level 4
Brute-Force	Generated 6 kept 4	Generated 15, kept 4	Generated 20, kept 4	Generated 20, kept 1
FKminus1F1	Generated 6, kept 4	Generated 6, kept 4	Generated 4, kept 1	-
FKminus1FKminus1	Generated 6, kept 4	Generated 6, kept 4	Generated 1, kept 1	-

2.2 Big Data Set

Method	Level 1	Level 2	Level 3	Level 4
Brute-Force	Generated 6kept 4	Generated 15, kept 4	Generated 20, kept 4	Generated 20, kept 1
FKminus1F1	Generated 122, kept 17	Generated 136, kept 15	-	-
FKminus1FKminus1	Generated 122, kept 17	Generated 136, kept 15	-	-