Introduction to Algorithmic Data Analysis (JADe)
Autumn 2023
Exercise sheet #3: Frequent Itemset Mining – FP-growth algorithm
Due 25.09.2023 at 06:00 (Exercise session on 25.09.2023 at 08:00, self-marking by the end of the day)

Please carefully read and follow the general instructions regarding exercises. Failing to meet the requirements might lead to penalties. *https://elearn.uef.fi/mod/page/view.php?id=293750*

If you suspect that something is wrong with some exercise question, please contact the lecturer.

If you face persistent issues while working on an exercise, do ask for help, e.g. during a course meeting or by contacting the lecturer via email.

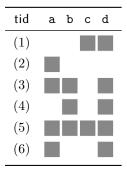Consider the dataset containing four items and six transactions represented as a binary matrix in Figure 1.

Figure 1: Transactional dataset, as a binary matrix

Problem 1 (FP-growth algorithm).

a) Construct the FP-tree representing the dataset. Assume an ordering of items by decreasing frequency. Show intermediate steps.

b) Run the FP-growth algorithm to mine all frequent itemsets at absolute minimum support of $\sigma = 2$. Assume an ordering of items by decreasing frequency. Show intermediate steps, i.e. intermediate conditional FP-trees, as well as the corresponding enumeration tree.