

## Data Mining proposal: Algorithm

Student name: Rodi Scheele

Student nr: 1013847

I would like to implement the Apriori Algorithm. The algorithm is used to find frequent item sets in a transactional database. The frequent item sets can be used to determine association rules that can show trends in the dataset.

The dataset I want to use is the 'Ta Feng' dataset. It contains 817741 transactions, belonging to 32266 users and 23812 items. A link to the dataset can be found here:

<https://stackoverflow.com/questions/25014904/download-link-for-ta-feng-grocery-dataset>

There is a lot of documentation on this algorithm available.

A few links with information that I would like to use are

<https://www.kdnuggets.com/2016/04/association-rules-apriori-algorithm-tutorial.html>

[http://www3.cs.stonybrook.edu/~cse634/lecture\\_notes/07apriori.pdf](http://www3.cs.stonybrook.edu/~cse634/lecture_notes/07apriori.pdf)

<http://software.ucv.ro/~cmihaescu/ro/teaching/AIR/docs/Lab8-Apriori.pdf>

I have found some improved Apriori Algorithms. After I have implemented the original algorithm, I would like to implement an improved version of it and compare the results with the original one.

<https://arxiv.org/ftp/arxiv/papers/1403/1403.3948.pdf>