Abstract

Frequent itemset mining leads to the discovery of associations among items in large transactional database. In this paper, two algorithms[7] of generating frequent itemsets are discussed: Apriori and FP-growth algorithm. In apriori algorithm candidates are generated and testing is done which is easy to implement but candidate generation and support counting is very expensive in this because database is checked many times. In the fp-growth, there is no candidate generation and requires only 2 passes over the database but in this the generation of fp-tree become very expansive to built and support is counted only when entire dataset is added to fp-tree. The comparison of these algorithms will tell which algorithm is better to perform.

References

1. J. Han, H. Pei, and Y. Yin. Mining Frequent Patterns without Candidate Generation. In: Proc. Conf. on the Management of Data (SIGMOD’00, Dallas, TX). ACM Press, New York, NY, USA 2000.
Comparative Study of Frequent Itemset Mining Algorithms Apriori and FP Growth


Index Terms

Computer Science

Artificial Intelligence

Keywords

Frequent itemset mining, Apriori, FP-Growth