Abstract

Mining of frequent patterns is a basic problem in data mining applications. Frequent Itemset Mining is considered to be an important research oriented task in data mining, due to its large applicability in real world applications. In this paper, a new Maximal Frequent Itemset mining algorithm with effective pruning mechanism is proposed. The proposed algorithm takes vertical tidset representation of the database and removes all the non-maximal frequent item-sets to get exact set of MFI directly. Pruning is done for both search space reduction and minimizing the number of frequency computations. It works efficiently when the number of item-sets and tid-sets are more. The proposed approach has been compared with Mafia algorithm for mushroom dataset and the results shows that the proposed algorithm performs effectively and generates frequent patterns faster. In order to understand the algorithm easily, an example is provided in detail.

References

- Roberto Bayardo, “Efficiently mining long patterns from databases”, in ACM
SIGMOD Conference 1998.
- K. Gouda and M. J. Zaki, "Efficiently Mining Maximal Frequent Itemsets", in Proc. of the IEEE Int. Conference on Data Mining, San Jose, 2001.
- Gosta Grahne and Jianfei Zhu, "Efficiently using prefix-trees in Mining Frequent Itemsets", in Proc. of the IEEE ICDM Workshop on Frequent Itemset Mining Implementations Melbourne, Florida, USA, November 19, 2003.
- J. Han, J. Pei, and Y. Yin, "Mining frequent patterns without candidate generation", In ACM SIGMOD Conf., May 2000.


Index Terms

Computer Science Data Mining
Keywords
Data Mining  Frequent Itemset Mining  Maximal Frequent Itemset Mining