Abstract

Data Mining is the multistage process of extraction of useful information from the large database. Association rule mining is one of the important techniques of data mining in which relationships among the items present in the transactions are discovered. There are different algorithms available in the field of data mining for association rule mining but most of them are time consuming hence the run time and memory overheads incurred is extremely high specially in the case of very large database. Sampling is one of the remarkable approach which can be used to speed up the process of association rule mining hence it is a approach to reduce the complexity of association rule mining technique to some extent but still consuming comparable time and memory. A progressive sampling based approach is a novel expert approach in the field of association rule mining to reduce the overheads of usual sampling based approaches. It is very effective in case of the large databases. In this paper, we have extended the Progressive sampling based approach presented by Umarani & Punithavalli,2009[22] and performed an extensive experimental analysis of the progressive sampling-based approach for the different Partitioned itemset 1/3,1/4,2/3,3/4 with the sample
An Improved Progressive Sampling based Approach for Association Rule Mining

dataset also in addition the performance of this Improved Progressive Sampling Based Approach is evaluated with the Progressive sampling based approach by Umarani & Punithavalli,2009[22]. The experimental results illustrate the complexity of an algorithm in terms of run time as well as the memory utilization. Complete implementation has been done in Java Jdk 6.1. and MySQL5.0 on the Sample dataset CompPeriPurchase.

References


12. Jiawei Han and Micheline Kamber “Data Mining –concept and techniques” Morgan Kaufmann Elsevier Science India (2002)


17. Srikant, R., Vu, Q. and Agarwal, R. “Mining association rules with item constraints”, In
the proceedings of 3rd Intl Conf on Knowledge Discovery and Data Mining, 1997.


**Index Terms**

Computer Science

Artificial Intelligence

**Keywords**

Association Rule Mining, Frequent Itemsets, Negative Border, Partitioned Itemsets.