Abstract

Mining frequent patterns in data is a useful requirement in several applications to guide future decisions. Association rule mining discovers interesting relationships among a large set of data items. Several association rule mining techniques exist, with the Apriori algorithm being common. Numerous algorithms have been proposed for efficient and fast association rule mining in databases, but these seem to only look at the data as a set of transactions, each transaction being a collection of items. The performance of the association rule technique mainly depends on the generation of candidate sets. In this paper we present a modified Apriori algorithm for discovering frequent items in data sets that are classified into categories, assuming that a transaction involves maximum one item being picked up from each category. Our specialized algorithm takes less time for processing on classified data sets by optimizing candidate generation. More importantly, the proposed method can be used for a more efficient mining of relational databases.

Reference

- M.S. Chen, J. Han and P.S. Yu. Data Mining: An overview from a database perspective. IEEE Transactions on Knowledge and Data Engineering, 1996.
- J. Han, J. Pei, and Y. Yin. Mining frequent patterns without candidate generation. ACM SIGMOD ICMD, 2000.
- M.H. Margahny and A.A. Mitwaly. Fast Algorithm for Mining Association Rules. AIML 05 Conf, Egypt.

Index Terms

Computer Science
Pattern Recognition
Keywords

Data mining  association rule

Apriori algorithm
transactions
frequent items
itemsets