Abstract
Apriori algorithm is a classic algorithm for frequent item set mining and association rule learning over transactional databases. The algorithm determines frequent item sets, which in turn can be used to determine association rules. These rules indicate the general trends in the database. Genetic algorithm is a search heuristic that mimics the process of natural selection using a greedy approach. This heuristic is routinely used to generate useful solutions for optimization and search problems. In this paper, we apply genetic algorithm to optimize the frequent item sets generated by Apriori algorithm and identify all possible significant association rules by analyzing the working of the algorithm on real data sets.

References

- Article titled &quot;How Amazon is leveraging Big Data&quot;, http://www.bigdata-startups.com/BigData-startup/amazon-leveraging-big-data/
- Anandhavalli M., Suraj Kumar Sudhanshu, Ayush Kumar and Ghose M. K., Optimized association rule mining using genetic algorithm, Advances in Information Mining, ISSN: 0975-3265, Volume 1, Issue 19 September 2011
- David Beasley et al., &quot;An Overview of Genetic Algorithms: Part 1, Fundamentals&quot;, University Computing, 1993
- Dataset from https://wiki.csc.calpoly.edu/datasets/attachment/wiki/apriori/apriori.zip
- Pang-Ning-Tan, Vipin Kumar, Michael Steinbach, (2007) Introduction to Data Mining

Index Terms

Computer Science
Artificial Intelligence

Keywords

Apriori  Genetic  Optimization  Transaction  Association Rule Mining