ISSN: 2348 - 2117

International Journal of Engineering Technology and Computer Research (IJETCR)
Available Online at www.ijetcr.org
Volume 4; Issue 2; March-April-2016; Page No. 43-46

IMPLEMENTING DCAR ALGORITHM IN SHOPPING COMPLEX DATASET

Dr. K.Kavitha

Assistant Professor

Department of Computer Science

Mother Teresa Women's University, Kodaikanal

kavitha.urc@gmail.com

Abstract

Data Mining has widely accepted a key technology for enterprises to improve their abilities in data analysis, decision making and automatic extraction of knowledge from data. Existing methods are subject to constraint that attaining sufficient level of domain knowledge for use in enterprise planning. Prediction in financial domain is absolutely difficult as well as predicting multi-dimensional schema. Existing DRCA algorithm is implemented for classifying association rules and generating prediction accuracy for Bank Loan. This paper implements the same algorithm in shopping complex dataset and reviewed experimental analysis.

Introduction

A huge range of data mining techniques has made significant improvements to the field of knowledge discovery in various domains. For example in the banking sector, these methodologies are used for loan payment prevision, classification of customers for targeted marketing, customer credit policy analysis, detection of money laundering schemes and other financial crimes. The banking database should control credit management thoroughly. Loan sanctioning requires the usage of huge data and significant processing time. To sanction the loan to customers, the bank needs to take some kind of precautions such as performance of the customer firm by analyzing the previous year's financial statements. The major tool used in multidimensional analysis in a data warehousing is the use of data aggregation and exploratory techniques that forms a part of (OLAP). The traditional OLAP technique is limited to detect hidden association between the items side in a data warehouse. So a lot of research undergoes to extend the OLAP technique to anticipate the future events. In this paper the capability of OLAP is extended to detect the hidden association and forecast the future events based on the historical data driven from the multidimensional schema.

The rest of the paper is organized as follows. Section II presents a description about the previous research which is relevant to the design and analysis of multidimensional schema. Section III describes about the existing model DDAR, DCAR, CMDC and its limitations. Section IV concludes this research work.

RELATED WORK:

Herawan and Deris presented an alternative approach for mining regular association rules and maximal association rules from transactional datasets. The transactional dataset was transformed into a Boolean-valued information system. The notions of regular and maximal association rules between two sets were defined. Also, support, confidence, maximal support, maximal confidences were defined using soft set theory [6]. *Kumar and A. Chadha* presented a case study of a university that hopes to improve the quality of education. So that association rule discovery was used to predict the accurate results [7]. *Romero, et al* explored the extraction of rare association rule from a Moodle system. Some relevant results were obtained and compared with the rare association rule mining algorithm [8]. *Zhu and Li* presented an association rule mining algorithm Ex- Apriori which was based on the predicate path graph. The algorithm can produce the predicate path graph by scanning database only once and dig out the frequent pattern based on the frequent predicate path graph. It avoids the shortcoming of scanning database many times [9].

Manda et al proposed an approach called Multi-ontology data mining at All Levels (MOAL) which uses

Corresponding author: Dr. K.Kavitha

the structure and relationships of the Gene Ontology (GO) to mine multi-ontology multi-level association rules. In this paper, two interesting measures was introduced i.e. Multi-ontology Support (MOSupport) and Multi-ontology Confidence (MO Confidence) customized to evaluate multi-ontology multi-level association rules.

Proposed Methodology

Data mining method can predict information that many traditional business analysis fail to deliver. Applications of Data Mining techniques yield the value of data warehouse by converting large volume of data into valuable for future business enhancement. Management Information System empowers the right people by providing the specific needed information. Association Rule Discovery has tried in this work to improvise the shopping complex Business strategy. Collected the data's like Purchasing details, Billing and stock details. Applying Association Rule method finds the similarity of the purchase & sales.

In this research work, lots of information are collected from the customer and vendor of shopping mall. Most frequently occurred query is

- Which products are most liked by Customer?
- Which combinations are most purchased by customers?

To identify this solution, Research has to undertaken this application to improve the business strategy by using DCAR algorithm. In appendix, DCAR algorithm is applied with the collected dataset and predicts some solutions based on the purchase and billing details at profit point of sale.

DCAR Algorithm

Multidimensional Schemas are generated based on the combination of hierarchical Clustering and Multi-dimensional scaling techniques. It identifies the significant numeric variables that define the cluster tools. Principal Component Analysis is used to rank the facts and dimensions. Attributes are ranked based on their nominal and numerical category. Data cube is constructed based on highest ranked dimensions and facts in the multi-dimensional schema. Based on the Rule importance from multi-dimensional schema and association rules are generated. The following figure shows the methodology of DCAR [12].

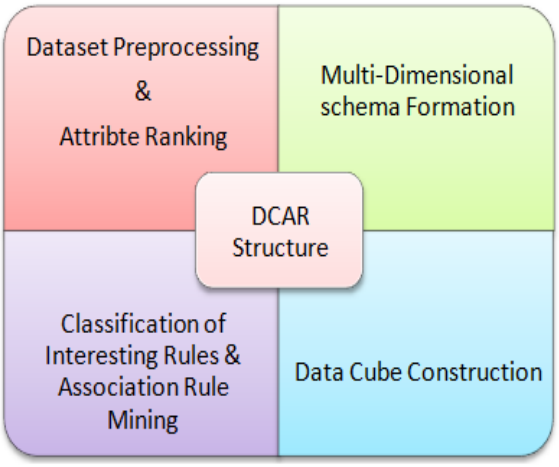

Figure 1:

Multi-Dimensional schema is formed by treating nominal variables as dimensions and numeric variables as facts. Information grouping defines the dimensional hierarchy or dimensional levels. Multiple fact tables shares the dimension tables. Multidimensional schema is generated from shopping complex dataset is given below.

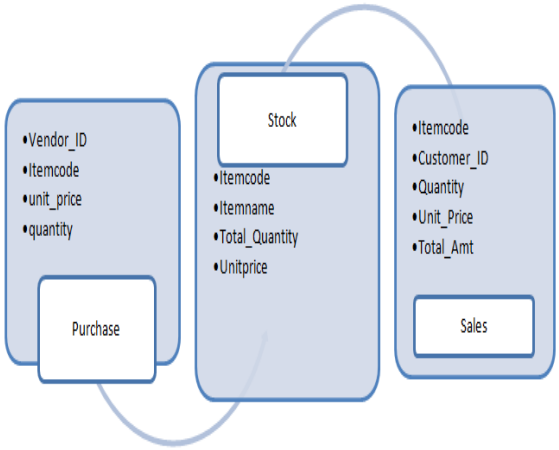

Figure 2:

Facts table represent stock details and Dimensions represents Sales, Purchase, Vendor and Customer details.

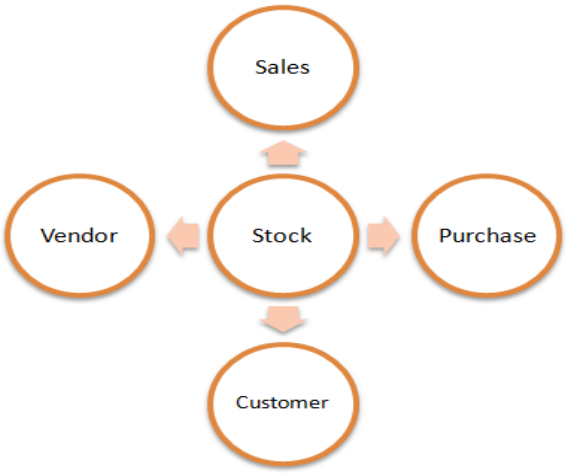

Figure 3:

The above multi-dimensional schema contains all the dimension and facts in the shopping complex dataset. It is used to construct informative data cubes. Highly ranked facts and dimensions are used to construct data cube. It helps the Consumer to show the rank positions. It stores combined measures of the dimension values in a multi-dimensional space. OLAP operations are used to explore the data in to meaningful information. Here facts and dimensions are separated based on the schema formation. In shopping complex database, Sales, purchase, customer and vendor are dimensional tables and stock kept in fact table.

Most frequent rules are extracted based on the user specified constraint. It finds the interesting relationships and associations among the attribute set effectively. Interested rules are predicted based on the constraint Weight which is to evaluated by support and confidence.

Discovered Rules are classified into highly interested, medium interested and low interested rules. To classify the rules, mean values are fixed based on highly ranked rules dimensions.
Discovered Rules are

Table 1:

Itemset	Weight
Shampoo	0.98
Conditioner	0.80
Oil	0.88
Detergent	0.70
Lotion	0.76
Shampoo, Lotion	0.86
Lotion, Oil	0.76
Soap, Oil	0.70
Shampoo, Conditioner	0.88
Lotion, Conditioner	0.80
Conditioner, Lotion, Soap	0.85
Lotion, Oil, Detergent	0.60
Soap, Oil, Lotion	0.56
Shampoo, Conditioner, Comb	0.55

Based on the constraint value such as weight, Highly Ranked Rules are classified separately as shown below

Table 2:

Itemset	Weight
Shampoo	0.98
Conditioner	0.80
Oil	0.88
Shampoo, Conditioner	0.88
Shampoo, Lotion	0.86
Lotion, Conditioner	0.80
Conditioner, Lotion, Soap	0.85

Medium Ranked Rules are

Table 3:

Itemset	Weight
Detergent	0.70
Lotion	0.76
Lotion, Oil	0.76
Soap, Oil	0.70

Low ranked rules are

Table 4:

Itemset	Weight
Lotion, Oil, Detergent	0.60
Soap, Oil, Lotion	0.56
Shampoo, Conditioner, Comb	0.55

Conclusion

In this paper, DCAR algorithm is implemented in the shopping complex dataset for developing the business strategy. This method will help the shopping mall to enhance the business strategy. As initiation process, Association Rule Mining and Multi-dimensional classification is applied later this application will be implemented in Clustering strategy.

References

1. J. K. Chiang and H. Sheng-Yin, "Multidimensional data mining for healthcare service portfolio management," in Computer Medical Applications (ICCMA), 2013 International Conference on, 2013, pp. 1-8.
2. P. Allard, S. Ferré, and O. Ridoux, "Discovering Functional Dependencies and Association Rules by Navigating in a Lattice of OLAP Views," in CLA, 2010, pp. 199-210.
3. W. Moudani, M. Hussein, M. Moukhtar, and F. Mora-Camino, "An intelligent approach to improve the performance of a data warehouse cache based on association rules," Journal of Information and Optimization Sciences, vol. 33, pp. 601-621, 2012.
4. M. Usman and S. Asghar, "An Architecture for Integrated Online Analytical Mining," Journal of Emerging Technologies in Web Intelligence, vol. 3, pp. 74-99, 2011.
5. H. Uğuz, "A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm," Knowledge-Based Systems, vol. 24, pp. 1024-1032, 2011.
6. T. Herawan and M. M. Deris, "A soft set approach for association rules mining," Knowledge-Based Systems, vol. 24, pp. 186-195, 2011.
7. V. Kumar and A. Chadha, "Mining Association Rules in Student's Assessment Data," International Journal of Computer Science Issues, vol. 9, pp. 211-216, 2012.
8. C. Romero, J. R. Romero, J. M. Luna, and S. Ventura, "Mining Rare Association Rules from e-Learning Data," in EDM, 2010, pp. 171-180.
9. H. Zhu and Q. Li, "An Algorithm Based on Predicate Path Graph for Mining Multidimensional Association Rules," in Proceedings of the 2012 International Conference on Information Technology and Software Engineering, 2013, pp. 783-791.
10. W. Abdelbaki, S. B. Yahia, and R. B. Messaoud, "NAP-SC: A Neural Approach for Prediction over Sparse Cubes," in Advanced Data Mining and Applications, ed: Springer, 2012, pp. 340-352.
11. J. Nahar, T. Imam, K. S. Tickle, and Y.-P. P. Chen, "Association rule mining to detect factors which contribute to heart disease in males and females," Expert Systems with Applications, 2012.
12. P. Manda, F. McCarthy, and S. M. Bridges, "Interestingness measures and strategies for mining multi-ontology multi-level association rules from gene ontology annotations for the discovery of new GO relationships," Journal of biomedical informatics, 2013.
13. H. R. Qodmanan, M. Nasiri, and B. Minaei-Bidgoli, "Multi objective association rule mining with genetic algorithm without specifying minimum support and minimum confidence," Expert Systems with applications, vol. 38, pp. 288-298, 2011.
14. C.-H. Weng and Y.-L. Chen, "Mining fuzzy association rules from uncertain data," Knowledge and Information Systems, vol. 23, pp. 129-152, 2010.
15. N. Zbidi, S. Faiz, and M. Limam, "On mining summaries by objective measures of interestingness," Machine learning, vol. 62, pp. 175-198, 2006.
16. K Kala Dr. E.Ramaraj "DCAR: A Novel Approach for Datacubes Association Rule Algorithm in Multidimensional Schema", CiiT International Journal of Data Mining and Knowledge Engineering, Vol 6, No 07, August 2014