

# A review of association rule mining using data mining

Surender singh | Yadwinder Singh | Yogesh Kumar

Department of CSE, BGIET, Sangru, Punjab, India

## To Cite this Article

Surender singh, Yadwinder Singh and Yogesh Kumar, "A review of association rule mining using data mining", *International Journal for Modern Trends in Science and Technology*, Vol. 04, Issue 03, March 2018, pp.-109-112.

DOI: <https://doi.org/10.46501/IJMTST040317>

## ABSTRACT

Association rule mining is an active data mining research area and most ARM algorithms cater to a centralized environment. Centralized data mining to discover useful patterns in distributed databases isn't always feasible because merging data sets from different sites incurs huge network communication costs. In this paper, an Improved algorithm based on good performance level for data mining is being proposed. In local sites, it runs the application based on the improved apriori algorithm, which is used to calculate support and confidence. now using apriori algorithm we add one more thing that should used for the stock management at the shop. it also tells about the how much frequency of item that should in the shop. So with help of New Apriori algorithm. The shopkeeper has less stress. They can easily maintain the stock of item. Now they also know well about how much one item is depend on another item. Their work is easy because they order items according to the sail of item. They have no need to reserve extra stock at shop. At last we get a benefit of each transaction. All of benefit is recorded using arm.

## Keywords

Data mining, frequent aided datasets, association rule, apriori algorithm.

Copyright © 2018 International Journal for Modern Trends in Science and Technology  
All rights reserved.

## I. INTRODUCTION

Data mining is an interdisciplinary research territory crossing several trains, for example, database frameworks; machine learning, wise data frameworks, measurements, and master frameworks. Data mining has formed into a key and element area of research by virtue of speculative troubles and practical applications related with the issue of discovering intriguing and already cloud gaining from boundless veritable databases. Various parts of data mining have been inquired about in a couple related fields. It has been get-together gigantic measures of information. Shockingly, these monstrous

aggregations of data set away on one of kind structures immediately moved toward getting to be plainly overwhelming. This basic pandemonium has incited the making of composed databases and database organization systems (DBMS). The viable database organization structures have been basic assets for organization of an inconceivable corpus of data and especially for practical and profitable recuperation of particular information from a broad social event at whatever point required.

## 1. Techniques of Data Mining:

### 1.1 Association

Association is one of the best known information mining procedures. In Association rule

find dependency of one item on other item. Here we know about the frequency of item using support and confidence

### 1.2 Classification

Classification is a consummate information removal process in vision of mechanism knowledge. Essentially arrangement is utilized to characterize everything in an arrangement of information into one of predefined set of classes or gatherings. organize plan make use of statistical scheme. In arrangement, we make the product that can figure out how to order the information things into gatherings. For instance, we can apply applicant teacher who are working in school. We can classify them subject wise teacher

### 1.3. Clustering

Clustering is an information mining system that makes significant or valuable group of articles that have comparative trademark utilizing programmed method. In this we make a group of predefined item like in library all books different category so we make a cluster of book according to syllabus

### 1.4 Prediction

The forecast examination strategy can be utilized as a part of offer to foresee benefit for the future on the off chance that we consider deal is a free factor, benefit could be a needy variable. At that point in light of the recorded deal and benefit information, we can draw a fitted relapse bend that is utilized revenue driven expectation.

### 1.5 Sequential Patterns

***Sequential examples investigation in one of information mining system that looks to find comparable examples in information exchange over a business period. The reveal examples are utilized for further business investigation to perceive connections among information.***

## 2. Development and Comparison of Novel Algorithm in ARM

Affiliation rules mirror the internal relationship of information. Finding these affiliations is advantageous for the right and fitting choice settled on by leaders. Affiliation control is a vital subject of information mining study. The affiliation rules give a successful intends to locate the potential connection between the information, mirroring an inherent relationship between the information. A few headways have been made for mining the proficient incessant examples from which the affiliation standards can be created. These strategies have certain downsides. There is no

technique which includes the benefit related with the things. To bolster this theory, a strategy is presented which mine the incessant item sets and assess the measure of benefit these successive itemsets are giving with the assistance of benefit measuring component. The after effects of the beforehand accessible calculation and new framed calculation are contrasted with assess the consequences of the proposed technique. The proposed method is validated via a series of experimentation conducted with many transactional datasets.

### 3. Association Rule Mining:-

Association rule mining, one of the most important and well researched techniques of data mining. It intends to remove intriguing successive examples, affiliations or easygoing structures among sets of things in the exchange databases or other information archives. Affiliation lead learning is a prominent technique for finding intriguing relations between factors in expansive databases. Affiliation control can be spoken to as

$$X \rightarrow Y$$

The itemset (set of items) X and Y are called "ANTECEDENT" and "CONSEQUENT" of the rule respectively...

#### USEFUL CONCEPTS:

- An itemset is a set of items.

E.g.,  $X = \{\text{milk, bread, butter}\}$  is an itemset.

- A  $k$ -itemset is an itemset with  $k$  items.

E.g.,  $\{\text{milk, bread, butter}\}$  is a 3-itemset

- SUPPORT:  $\text{SUPP}(X)$  of an itemset X is defined as the proportion of transactions in dataset which contain itemset and the total number of transactions.
- CONFIDENCE is defined as the ratio of number of transactions that include all the items in the consequent as well as the antecedent to the number of transactions that include all the items in antecedent.

$$\text{CON}(X \rightarrow Y) = \frac{\text{SUPP}(XY)}{\text{SUPP}(X)}$$

Example- this fig 3.1 (transaction of items) shows the total 12 transaction and 3 item are selected to be count the support and confidence of item.



| TID | ITEM1 | ITEM2 | ITEM3 |
|-----|-------|-------|-------|
| 1   | 0     | 0     | 4     |
| 2   | 0     | 7     | 0     |
| 3   | 0     | 2     | 9     |
| 4   | 5     | 0     | 0     |
| 5   | 11    | 0     | 13    |
| 6   | 6     | 9     | 8     |
| 7   | 9     | 3     | 6     |
| 8   | 14    | 16    | 12    |
| 9   | 0     | 5     | 7     |
| 10  | 2     | 0     | 3     |
| 11  | 7     | 10    | 5     |
| 12  | 0     | 9     | 0     |

fig 3.1 transaction of items

Support of item 1 =  $7/12$

Supp = 0.58

Confidence of item 1 and 2 =  $0.33/0.58$

Conf = 0.57

This table is used for taking further result

#### 4. Result

The result of this shown at last as shown in the fig 4.1 profit. Now we get result in the form of profit. It is also depend upon the support and confidence of a item. The result is displayed using the vb.net framework. According to the choice we can select the min support and confidence.

| SUPPORT                       | CONFIDENCE                    | PROFIT                 |
|-------------------------------|-------------------------------|------------------------|
| Support of Item(s)-1 = 0.58   | Confidence of Item(12) = 0.57 | Profit of (1) = 972    |
| Support of Item(s)-2 = 0.67   | Confidence of Item(13) = 0.86 | Profit of (2) = 2623   |
| Support of Item(s)-3 = 0.75   | Confidence of Item(23) = 0.75 | Profit of (3) = 2077   |
| Support of Item(s)-12 = 0.33  |                               | Profit of (12) = 2282  |
| Support of Item(s)-13 = 0.5   |                               | Profit of (13) = 2339  |
| Support of Item(s)-23 = 0.5   |                               | Profit of (23) = 3392  |
| Support of Item(s)-123 = 0.33 |                               | Profit of (123) = 3243 |

Fig 4.1 profit

#### 5. Need of algorithm:-

The proposed research work is based on the formation of new algorithm which is based upon the primitive Apriori algorithm that will augment the attributes and efficiency and reduce time by making a model of prototype which will be beneficial in overcoming the shortcomings of apriori algorithm. We theoretically and experimentally analyze the apriori Algorithm which is the most established algorithm for frequent itemset mining. The work is focused on apriori Algorithm.

#### 6. Conclusion:-

We also develop a new algorithm in which stock management is very useful it tells us about how much frequency of item left and what frequency is depend on another using this paper every shopkeeper does not need of any extra stock all of the record of every item have available in this program. So we can say that it is very useful for the shopkeeper and this is one of the new thing that is discovered first time only because of the demand of the shopkeeper. It is possible only using association techniques most of the things are related with each other so here this rule is used.

#### REFERENCES

- [1]. Arti Rathore, Ajaysingh Dhabariya and Chintan Thacker: "A Review on Association Rule Mining and Improved Apriori Algorithms", International Journal of Scientific Research in Computer Science, Vol. 1, No. 1, pp. 1-5, (September, 2013).
- [2]. Du Ping, Gao Yongping: "A New Improvement of Apriori Algorithm for Mining Association Rules", International Conference on Computer Application and System Modeling, ISBN: 978-1-4244-7237-6, pp. V2: 529-532, (October, 2010).
- [3]. Goswami D.N., Chaturvedi Anshu, Raghuvanshi C.S: "An Algorithm for Frequent Pattern Mining Based On Apriori", International Journal on Computer Science and Engineering (ISSN: 0975-3397), Vol. 2, No. 4, pp. 942-947 (August, 2010).
- [4]. A.M.J. Md. Zubair Rahman and P. Balasubramanie: "Weighted Support Association Rule Mining using Closed Itemset Lattices in Parallel", International Journal of Computer Science and Network Security (ISSN : 1738-7906), ol. 9 No. 3 pp. 247-253, (March, 2009).
- [5]. Chun-Jung Chu, Vincent S. Tseng and Tyne Liang: "An efficient algorithm for mining high utility itemsets with negative item values in large databases", Applied Mathematics and Computation (ISSN: 0096-3003), Vol. 215, No. 2, pp.767-778, (September, 2009).
- [6]. HAN Feng, ZHANG Shu-mao and DU Ying-shuang: "The analysis and improvement of Apriori algorithm", Journal of Communication and Computer (ISSN: 1548-7709), Vol. 5, No. 9, pp. 12-18, (September, 2008).
- [7]. Wei Zhang, Hongzhi Liao and Na Zhao: "Research on the FP Growth Algorithm about Association Rule Mining", International Seminar on Business and Information Management (ISBN: 978-0-7695-3560-9), pp. 315-318, (December, 2008).
- [8]. Guangzhu Yu, Shihuang Shao and Xianhui Zeng: "Mining Long High Utility Itemsets in Transaction Databases", Mining Long High Utility Itemsets in Transaction Databases (ISSN: 1790-0832), Vol. 2, No. 5, pp. 202-210, (February, 2008).
- [9]. Jianying Hu and Aleksandra Mojsilovic: "High-utility pattern mining: A method for discovery of high-utility item sets", Pattern Recognition, Vol. 40, No. 11, pp. 3317-3324, (November, 2007).
- [10]. Sotiris Kotsiantis, Dimitris Kanellopoulos: "Association Rules Mining: A Recent Overview", GESTS International Transactions on Computer Science and Engineering, Vol. 32, No. 1, pp. 71-82, (January, 2006).
- [11]. Maria N. Moreno, Saddys Segre and Vivian F. Lopez: "Association Rules: Problems, solutions and new applications", TAMIDA2005 (ISBN: 84-9732-449-8), pp. 317-323, (June, 2005).
- [12]. Oded Z. Maimon, Lior Rokach: "Decomposition Methodology for Knowledge Discovery and Data Mining:

Theory and Applications", World Scientific Publishing Company, ISBN: 9789812560797, (May 2005).

- [13]. Unil Yuna: "Efficient mining of weighted interesting patterns with a strong weight and/or support affinity", Information Sciences, Vol. 177, No. 17, pp. 3477-3499, (September 2007).

