

# An Overview of Classification Rule and Association Rule Mining

B. Srinivas<sup>1</sup>, Gadde Ramesh<sup>2</sup>, Shoban Babu Sriramoju<sup>3</sup>

<sup>1</sup>Associate Professor, Department of Computer Science, Madda Walabu University, Ethiopia

<sup>2</sup>Assistant Professor, School of Electrical and Computer Engineering, Ethiopian Institute of Technology, Ethiopia

<sup>3</sup>Professor, Department of Computer Science and Engineering, SR Engineering College, Telangana, India

## ABSTRACT

In this study, we focused on how data mining can be applied in Market Basket Analysis to identify new trend and purchasing patterns of customers. Data Mining is the process of extracting useful information from a large dataset. We are here using Association Rule to identify the relationship between different products and Classification Rule on the consumers to distinguish them on the basis of pre-defined parameters. The most important information discovered from the combination of these two algorithms supports the retailers in decision making and hence increase their sales.

**Keywords :** Classification Rule, Association Rule, Clustering, Outlier Analysis

## I. INTRODUCTION

In this day and age large measure of data is generated every day and this data is maintained in database in different fields, for example, healthcare, education, market basket analysis, etc. With this increasing data size, there is a need to understand large and complex data and reach necessary determinations. The process of extracting necessary data from large pre-existing databases is called Data Mining. It is very troublesome for the neighborhood retailers to pull in customers, so it is their need to understand the shopping trends of the customers.

Numerous consumers prefer online shopping. With the development of the e-commerce websites, retailers tend to neglect to pull in more and more consumers. This problem can be eliminated by applying data mining techniques to analyze new patterns and trends. The data mining techniques are

applied to the gathered data associated to customer behavior pattern, with the goal that retailers will be able to know the new patterns and trends.

### A. STEPS INVOLVED IN DATA MINING

- a. Identifying the source information
- b. Picking the data points that need to be analyzed
- c. Extracting the relevant information from the data
- d. Identifying the key values from the extracted data set
- e. Interpreting and reporting the results

## II. DATA MINING TECHNIQUES

### A. Association Rule

This rule is incorporated for establishing the relationship between different objects that exist in

the market. [1] This rule is very helpful in Market Basket Analysis.

### B. Classification

This rule is used to arrange a data item into predefined classes. For example we can use this rule to differentiate autos into different categories (KUV, Sedan, and SUV). Same principles can be applied to consumers, for e.g. by recognizing them into income, age and social gathering.

### C. Clustering

With this Clustering technique, data is organized and classified into meaningful subgroups or clusters.

### D. Prediction

This technique is used to predict new information from a set of existing data. For example Sales in future week can be predicted using this rule.

### E. Outlier Analysis

This technique identifies and explains exceptions. For example in Market Basket Data Analysis, Outlier can be some transactions which happen uncommon.

## III. EXISTING ALGORITHM

There are so many algorithms used in Market Basket Analysis for identifying the changes in purchasing trends of customers. Some algorithms are only focused on the customers while others are focused only on the products.

## IV. PROPOSED METHOD

In this paper, we combine Association Rule and Classification Rule which will focus on both the products and the customer's. Association Rule will be implied on the products and Classification Rule will be implied on the customers. So this will be more affectionate for the retailers to determine the purchasing trends of customers.

## V. ASSOCIATION RULE

This rule is specifically used for making the relationship between two or more items. This rule can be used for tracing customers buying habits. We can observe that when a consumer always purchases shoes then he purchases socks also and therefore we can suggest that in future when he purchases shoes he might also want to purchase socks.

### Example:

Table 1. An example of Market Basket Transactions.

Transaction ID (TID)	Items
1.	Butter, Cheese, Burger
2.	Milk, Cheese, Butter
3.	Butter, Milk

Here we can identify the relationship between the Butter and Milk. Here the Association rule shows us that whenever a customer buys Milk he always buys butter along with it.

Milk->Butter

This rule is helpful for retailers to identify the obtaining patterns of the customers. And they will hence come to think about what the customer needs. Understanding Association rule with example

The dataset below contains the products purchased from a retailer.

Tr.ID	Product 1	Product 2	Product 3
1	Bread	Butter	Jam
2	Bread	Butter	Jam
3	Bread	Butter	Jam
4	Bread	Butter	
5	Bread	Butter	

Tr.ID-Transaction ID

For the above dataset we can establish the following relations between the different items :

**Rule no 1:** IF Bread is bought, THEN butter is also bought.

**Rule no 2:** IF butter is bought, THEN Bread is also bought.

**Rule no 3:** IF Bread and butter is brought, THEN Jam is also brought in 60% of the transactions.

“IF THEN” format is followed in Association Rule.

These terminologies that are defined are adapted from Data Mining for Business Intelligence, by GalitShmueli and others[4].

**Support:** This rule signifies the effect in terms of overall size. In the event that exclusive a few number of exchanges are affected, the rule might be of less use. For e.g, the help of "IF Bread and Butter THEN Jam" is 3/5 i.e. 60% of the aggregate exchanges.

**Confidence:** This determines functional use of the rule. Transactions that are having confidence greater than 50% are selected. For e.g., confidence of Bread, butter and Jam given can be written as:

The exchanges containing Bread and Butter (Antecedent) and Jam (Consequent) are 3

The exchanges containing just Bread and Butter (Antecedent) are 5.

$P(\text{Butter and Bread and Jam})/P(\text{Butter and Bread}) = 3/5 = 60\%$

Thusly, we can make a conclusion that the association rule is having a confidence of 60%.

## VI. CLASSIFICATION RULE

Here classification rule mining can be used for differentiating customers or items on the basis of different parameters. Here the Classification rule can be used to differentiate the customers by frequency of visits, age, income, marital status, etc.

Classification rule helps the retailers to take decisions easily.

Age	Frequency	Marital Status	Frequency
Below 20	30	Single	45
20-40	50	Married	45
Above 40	40		
Income	Frequency	Gender	Frequency
<1 Lakh	20	Male	70
1Lakh-5 Lakh	45	Female	50
>5 Lakh	25		

Retailers will have the knowledge about the customers according to these defined parameters, in this case age, marital status, income, gender.

Suppose in an area if there are more number of students (age <20) so the retailer will focus more towards the needs of the student like stationary items and sports items.

## VII. CONCLUSION

Data Mining has played a very important role in Market Analysis and various other fields. The most important point to succeed in a marketing strategy is to create an accurate customer analysis[2]. The motivation for applying data mining approach on Market Basket Analysis is to learn about buying patterns and retailers can use this information so more no of consumers are attracted towards them.

## VIII. REFERENCES

- [1]. ManpreetKaur ,Shivani Kang\*,Market Basket Analysis: Identify the changing trends of market data using association rule mining,
- [2]. NeseAcar, BülentÇizmeçi, Factors Influencing Customer’s Choice of Technology Retailers: An Application In Kayseri (Turkey),
- [3]. NeeshaJothiNur’Aini,AbdulRashidWahidahHusain, Data Mining in Healthcare – A Review,

- <https://www.analyticsvidhya.com/blog/2014/08/effective-cross-selling-market-basket-analysis/>
- [4]. Shoban Babu Sriramoju, "Analysis and Comparison of Anonymous Techniques for Privacy Preserving in Big Data" in "International Journal of Advanced Research in Computer and Communication Engineering", Vol 6, Issue 12, December 2017, DOI 10.17148/IJARCCCE.2017.61212 [ ISSN(online) : 2278-1021, ISSN(print) : 2319-5940 ]
- [5]. Shoban Babu Sriramoju, " Review on Big Data and Mining Algorithm" in "International Journal for Research in Applied Science and Engineering Technology", Volume-5, Issue-XI, November 2017, 1238-1243 [ ISSN : 2321-9653], [www.ijraset.com](http://www.ijraset.com)
- [6]. Shoban Babu Sriramoju, "OPPORTUNITIES AND SECURITY IMPLICATIONS OF BIG DATA MINING" in "International Journal of Research in Science and Engineering", Vol 3, Issue 6, Nov-Dec 2017 [ ISSN : 2394-8299 ].
- [7]. Guguloth Vijaya, A. Devaki, Dr. Shoban Babu Sriramoju, "A Framework for Solving Identity Disclosure Problem in Collaborative Data Publishing" in "International Journal of Research and Applications" (Apr-Jun © 2015 Transactions), Vol 2, Issue 6, 292-295 [1
- [8]. Shoban Babu Sriramoju, "Heat Diffusion Based Search for Experts on World Wide Web" in "International Journal of Science and Research", <https://www.ijsr.net/archive/v6i11/v6i11.php>, Volume 6, Issue 11, November 2017, 632 - 635, #ijsrnet
- [9]. Dr. Shoban Babu Sriramoju, Prof. Mangesh Ingle, Prof. Ashish Mahalle "Trust and Iterative Filtering Approaches for Secure Data Collection in Wireless Sensor Networks" in "International Journal of Research in Science and Engineering" Vol 3, Issue 4, July-August 2017 [ ISSN : 2394-8299 ].
- [10]. Dr. Shoban Babu, Prof. Mangesh Ingle, Prof. Ashish Mahalle, "HLA Based solution for Packet Loss Detection in Mobile Ad Hoc Networks" in "International Journal of Research in Science and Engineering" Vol 3, Issue 4, July-August 2017 [ ISSN : 2394-8299 ].
- [11]. Shoban Babu Sriramoju, "A Framework for Keyword Based Query and Response System for Web Based Expert Search" in "International Journal of Science and Research" Index Copernicus Value(2015):78.96 [ ISSN : 2319-7064 ].
- [12]. Sriramoju Ajay Babu, Dr. S. Shoban Babu, "Improving Quality of Content Based Image Retrieval with Graph Based Ranking" in "International Journal of Research and Applications" Vol 1, Issue 1, Jan-Mar 2014 [ ISSN : 2349-0020 ].
- [13]. Dr. Shoban Babu Sriramoju, Ramesh Gadde, "A Ranking Model Framework for Multiple Vertical Search Domains" in "International Journal of Research and Applications" Vol 1, Issue 1, Jan-Mar 2014 [ ISSN : 2349-0020 ].
- [14]. Mounika Reddy, Avula Deepak, Ekkati Kalyani Dharavath, Kranthi Gande, Shoban Sriramoju, "Risk-Aware Response Answer for Mitigating Painter Routing Attacks" in "International Journal of Information Technology and Management" Vol VI, Issue I, Feb 2014 [ ISSN : 2249-4510 ]
- [15]. Mounica Doosetty, Keerthi Kodakandla, Ashok R, Shoban Babu Sriramoju, "Extensive Secure Cloud Storage System Supporting Privacy-Preserving Public Auditing" in "International Journal of Information Technology and Management" Vol VI, Issue I, Feb 2012 [ ISSN : 2249-4510 ]
- [16]. Shoban Babu Sriramoju, "An Application for Annotating Web Search Results" in "International Journal of Innovative Research in Computer and Communication Engineering" Vol 2, Issue 3, March 2014 [ ISSN(online) : 2320-9801, ISSN(print) : 2320-9798 ]
- [17]. Shoban Babu Sriramoju, "Multi View Point Measure for Achieving Highest Intra-Cluster Similarity" in "International Journal of Innovative Research in Computer and

- Communication Engineering" Vol 2, Issue 3, March 2014 [ ISSN(online) : 2320-9801, ISSN(print) : 2320-9798 ]
- [18]. Shoban Babu Sriramoju, Madan Kumar Chandran, "UP-Growth Algorithms for Knowledge Discovery from Transactional Databases" in "International Journal of Advanced Research in Computer Science and Software Engineering", Vol 4, Issue 2, February 2014 [ ISSN : 2277 128X ]
- [19]. Shoban Babu Sriramoju, Azmera Chandu Naik, N.Samba Siva Rao, "Predicting The Misusability Of Data From Malicious Insiders" in "International Journal of Computer Engineering and Applications" Vol V, Issue II, February 2014 [ ISSN : 2321-3469 ]
- [20]. Ajay Babu Sriramoju, Dr. S. Shoban Babu, "Analysis on Image Compression Using Bit-Plane Separation Method" in "International Journal of Information Technology and Management", Vol VII, Issue X, November 2014 [ ISSN : 2249-4510 ]
- [21]. Shoban Babu Sriramoju, "Mining Big Sources Using Efficient Data Mining Algorithms" in "International Journal of Innovative Research in Computer and Communication Engineering" Vol 2, Issue 1, January 2014 [ ISSN(online) : 2320-9801, ISSN(print) : 2320-9798 ]
- [22]. Ajay Babu Sriramoju, Dr. S. Shoban Babu, "Study of Multiplexing Space and Focal Surfaces and Automultiscopic Displays for Image Processing" in "International Journal of Information Technology and Management" Vol V, Issue I, August 2013 [ ISSN : 2249-4510 ]
- [23]. Dr. Shoban Babu Sriramoju, "A Review on Processing Big Data" in "International Journal of Innovative Research in Computer and Communication Engineering" Vol-2, Issue-1, January 2014 [ ISSN(online) : 2320-9801, ISSN(print) : 2320-9798 ]
- [24]. Shoban Babu Sriramoju, Dr. Atul Kumar, "An Analysis around the study of Distributed Data Mining Method in the Grid Environment : Technique, Algorithms and Services" in "Journal of Advances in Science and Technology" Vol-IV, Issue No-VII, November 2012 [ ISSN : 2230-9659 ]
- [25]. Shoban Babu Sriramoju, Dr. Atul Kumar, "An Analysis on Effective, Precise and Privacy Preserving Data Mining Association Rules with Partitioning on Distributed Databases" in "International Journal of Information Technology and management" Vol-III, Issue-I, August 2012 [ ISSN : 2249-4510 ]
- [26]. Shoban Babu Sriramoju, Dr. Atul Kumar, "A Competent Strategy Regarding Relationship of Rule Mining on Distributed Database Algorithm" in "Journal of Advances in Science and Technology" Vol-II, Issue No-II, November 2011 [ ISSN : 2230-9659 ]
- [27]. Shoban Babu Sriramoju, Dr. Atul Kumar, "Allocated Greater Order Organization of Rule Mining utilizing Information Produced Through Textual facts" in "International Journal of Information Technology and management" Vol-I, Issue-I, August 2011 [ ISSN : 2249-4510 ]
- [28]. Ramesh Gadde, Namavaram Vijay, "A SURVEY ON EVOLUTION OF BIG DATA WITH HADOOP" in "International Journal of Research in Science and Engineering", Vol-3, Issue-6, Nov-Dec 2017, 92-99 [ ISSN : 2394-8299 ].
- [29]. Namavaram Vijay, S Ajay Babu, "Heat Exposure of Big Data Analytics in a Workflow Framework" in "International Journal of Science and Research", Volume 6, Issue 11, November 2017, 1578 - 1585, #ijsrnet
- [30]. Ajay Babu Sriramoju, Namavaram Vijay, Ramesh Gadde, "SKETCHING-BASED HIGH-PERFORMANCE BIG DATA PROCESSING ACCELERATOR" in "International Journal of Research in Science and Engineering", Vol-3, Issue-6, Nov-Dec 2017, 92-99 [ ISSN : 2394-8299 ].
- [31]. Namavaram Vijay, Ajay Babu Sriramoju, Ramesh Gadde, "Two Layered Privacy

Architecture for Big Data Framework" in "International Journal of Innovative Research in Computer and Communication Engineering" Vol 5, Issue 10, October 2017 [ ISSN(online) : 2320-9801, ISSN(print) : 2320-9798 ]

- [32]. Amitha Supriya. "Implementation of Image Processing System using Big Data in the Cloud Environment." International Journal for Scientific Research and Development 5.10 (2017): 211-217.
- [33]. SA Supriya. "A Survey Model of Big Data by Focusing on the Atmospheric Data Analysis." International Journal for Scientific Research and Development 5.10 (2017): 463-466.
- [34]. Siripuri Kiran, 'Decision Tree Analysis Tool with the Design Approach of Probability Density Function towards Uncertain Data Classification', International Journal of Scientific Research in Science and Technology(IJSRST), Print ISSN : 2395-6011, Online ISSN : 2395-602X, Volume 4 Issue 2, pp.829-831, January-February 2018. URL : <http://ijsrst.com/IJSRST1841198>
- [35]. Ajmera Rajesh, Siripuri Kiran, " Anomaly Detection Using Data Mining Techniques in Social Networking" in "International Journal for Research in Applied Science and Engineering Technology", Volume-6, Issue-II, February 2018, 1268-1272 [ ISSN : 2321-9653], [www.ijraset.com](http://www.ijraset.com)