International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 04 Issue: 11 | Nov -2017 www.irjet.net p-ISSN: 2395-0072

Reducing the gap between Consumer and Retailer using Association Rule Mining and Classification Rule: An Overview

Arpit Kalra[1], Harsh Mehta[2], Om laddha[3]

[1,2,3] *SVKM's NMIMS UNIVERSITY, Shirpur*

Abstract: *In this study, we focused on how data mining can be applied in Market Basket Analysis to identify new trend sand 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 valuable information discovered from collaboration of these two algorithms supports the retailers in decision making and hence increase their sales.*

I. INTRODUCTION

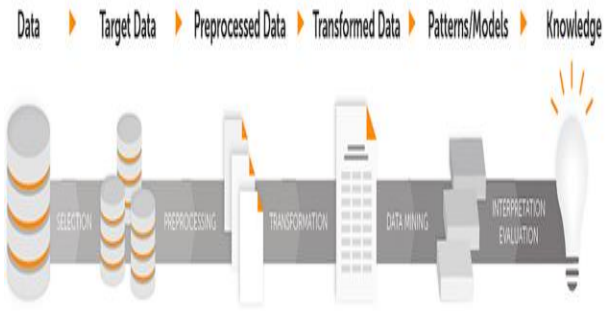
In today's world large amount of data is generated every day and this data is maintained in database in various fields such as healthcare, education, market basket analysis, etc. With this increasing data size, there is a need to understand large and complex data and draw necessary conclusions. The practice of extracting useful information from large pre-existing databases is called Data Mining. It is getting difficult for the local retailers to attract customers, so there is a need for the retailers to understand the shopping trends of the customers.

Many consumers prefer online shopping. With the growth of the e-commerce websites, retailers tend to fail to attract more and more consumers.

And this problem can be resolved by applying data mining techniques to analyze new patterns and trends. We will apply data mining techniques to the gathered data of the customer behavior pattern, so that retailers will 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 used to establish a relationship between different objects that exist in the market. [1] This rule is helpful in Market Basket Analysis.

B. Classification

This rule is used to classify a data item into predefined classes. Foreg, we can use this rule to simply differentiate cars into various categories. (Sedan, SUV)
Same principles can be applied to consumers, for e.g. by distinguishing them into income, age and social group.

C. Clustering

With this 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 e.g., in Market Basket Data Analysis, Outlier can be some transactions which happen uncommon.

III. EXISITING ALGORITHM

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

IV. PROPOSED METHOD

We are proposing in this paper to collaborate Association 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 effective for the retailers to determine the purchasing trends of customers.

V.ASSOCIATION RULE

Association rule is useful for making a simple relationship between two or more items. This rule can be used for tracing customers buying habits.

We might observe that a consumer always purchases socks when he purchases shoes, & therefore suggest that in future when he purchases shoes he might also want to purchase socks.

Example:

Table1. An example of Market Basket Transactions.

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

Here we identify an important relationship between Milk and butter. 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 purchasing patterns of the customers. And they will hence come to know about what the customer wants.
Understanding Association rule with example

This is a dataset containing items 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 following relations between the different items:

Rule no 1: IFBread is bought, THEN butter is also bought.

Rule no 2: IF butter is bought, THENBread is also bought.

Rule no 3: IFBread and butter is bought, THENJam is also bought in 60% of the transactions.
"IF THEN" format is followed in Association Rule.
The following terminologies defined are adapted from Data Mining for Business Intelligence, by GalitShmueli and others[4].

Frequent ProductSet :– Product set that is purchased most number of times. For e.g., in the above dataset, Bread&butterappeared in 100% of the purchases made.

Support :-This rule signifies the effect in terms of overall size. If only a fewnumber of transactions are affected, the rule may be of less use. For e.g, the support of "IF BreadandButter THEN Jam" is 3/5 i.e. 60% of the total transactions.

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

The transactions containingBread&Butter (Antecedent) and Jam (Consequent) are 3
The transactions containing only Bread&Butter (Antecedent) are 5.
P(Bread&ButterandJam)/P (BreadandButter) = 3/5 = 60%
Thus we can conclude that the association rule is having a confidence of 60%. More the confidence, Stronger the rule is.

VI. CLASSIFICATION RULE

We can use classification rule mining to differentiate customers or items on the basis of different parameters.
Here we are using Classification rule to classify the customers by age, income, frequency of visits, 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
1 Lakh-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]NeşeAcar, BülentÇizmeci, Factors Influencing Customer's Choice of Technology Retailers: An Application In Kayseri (Turkey),

[3]NeeshaJothiNur'Aini,AbdulRashidWahidahHusain, Data Mining in Healthcare – A Review,

[4]https://www.analyticsvidhya.com/blog/2014/08/effective-cross-selling-market-basket-analysis/