

Association Rule – Extracting Knowledge Using Market Basket Analysis

Summarizing the article.
Assignment 9

Tania Cajal (MBDS 2021/2022)

Raorane A.A.¹ Kulkarni R.V.² and Jitkar B.D.³

Organizations' ability to make decisions and understand user behaviour has become a critical and difficult problem to solve in order to maintain their competitive market position. Technological advancements have opened the way for speedier query processing and sub-second response times. Data mining technologies have become the most reliable weapon for analyzing large amounts of data and generating accurate conclusions. Data collected in large databases has become “Data mountains”, which are rarely visited. The goal is to evaluate a large quantity of data in order to exploit customer behavior and make the best decision possible, giving you a competitive advantage over your competitors. To demonstrate its superiority to traditional approaches, an experimental analysis was conducted utilizing association rules and Market Basket Analysis.

Knowledge discovery as a process consists of an iterative sequence of the following steps:

- i) Data Cleansing (to remove noise and inconsistent data).
- ii) Data integration (where multiple data sources may be combined).
- iii) Data Selection (Where data relevant to the analysis task are retrieved from the database).
- iv) Data transformation (Where data are transformed or consolidated into forms appropriate for mining by performing summary or aggregation operations for

instance.

v) Data mining (an essential process where intelligent methods are applied in order to extract data patterns).

vi) Pattern evolution (to identify the truly interesting patterns representing knowledge based on some interestingness measures).

vii) Knowledge presentation (Where visualization and knowledge representation techniques are used to present

Data Collection Method: From the supermarket called Shetkari Bazar in Kolhapur city in Maharashtra, India, the day today transactional data is gathered. The sale of products are the actual transactions made by the customer, which consist of various products. Each item has a Boolean variable representing by a Boolean vector of values assigned to these variables. The Boolean vector can be analyzed for buying patterns.

These patterns can be represented in the form of association rules. For example, the information that customers who purchase computers also tend to buy printer at the same time is represented in Association Rule in the article's example. Computer = Printer Support = 20%, Confidence = 80%

Market Basket Transaction of Market Basket Analysis is a data mining technique to derive association between data sets. We have categorical data of transaction records as input to the analysis and the output of the analysis are association rules as a new knowledge directly from stored data.

Market Basket Analysis:

i) Step by step computation of Market Basket Analysis is as follow.

ii) Generate all possible association rules.

iii) Compute the support and confidence of all possible association rules.

iv) Apply two threshold criteria: minimum support and minimum confidence to obtain the association rule.

Using this process, it is discovered that there is a connection between the products at the moment of purchase by the clients.

Furthermore, it has been discovered that this approach is best suited for regulating product placement on supermarket shelves. This strategy may result in a higher

profit for the seller. As a result, the Data Mining tool can be utilized to optimize the product placement strategy on the shelf by applying Data Mining techniques.