

Shantanu S. Mangak  
B.E B 41242 DMW  
LPII

## Assignment No. 3

Title: Apply A-priori algorithm to find frequently occurring items from given data & generate strong association rules using support & confidence thresholds.

Problem Definition: Market Basket Analysis.

Pre-requisite :-

Basic Concepts of ETL

S/W & H/W Requirements:-

Rapidminer, PIV, 2GB RAM, 500 GB HDD.

Learning Objectives:-

Model associations between products by determining sets of items frequently purchased together & building association rules to derive recommendations.

Outcomes :-

Create association rules which can be used for product recommendations depending on the confidences of the rules.

## Theory:-

- Association rule for mining:
  - i) Proposed by R. Agrawal & R. Srikant in 1994.
  - ii) Proposed by R. Agrawal — " —
  - iii) It is an important data mining model studied extensively by the database & data mining community.
  - iv) Assume all data are categorical. Initially used for market Basket Analysis to find how items purchased by customers are related.

## • A-priori algorithm:

The best known algorithm. Two steps:

- i) Find all items sets that have minimum support (frequent itemsets, also called large item sets).
- ii) It create Association rule with support & Confidence.

## • Market Basket Analysis using Rapid Miner:

Rapid Miner is a data science software platform developed by the company of the same name that provides an integrated environment for data preparation, ML, DL, text mining & predictive analytics. It is used for business & commercial applications as well as for research, education, training, rapid

prototyping & application development & supports all steps of the ML process including data preparation.

- Market Basket Analysis:

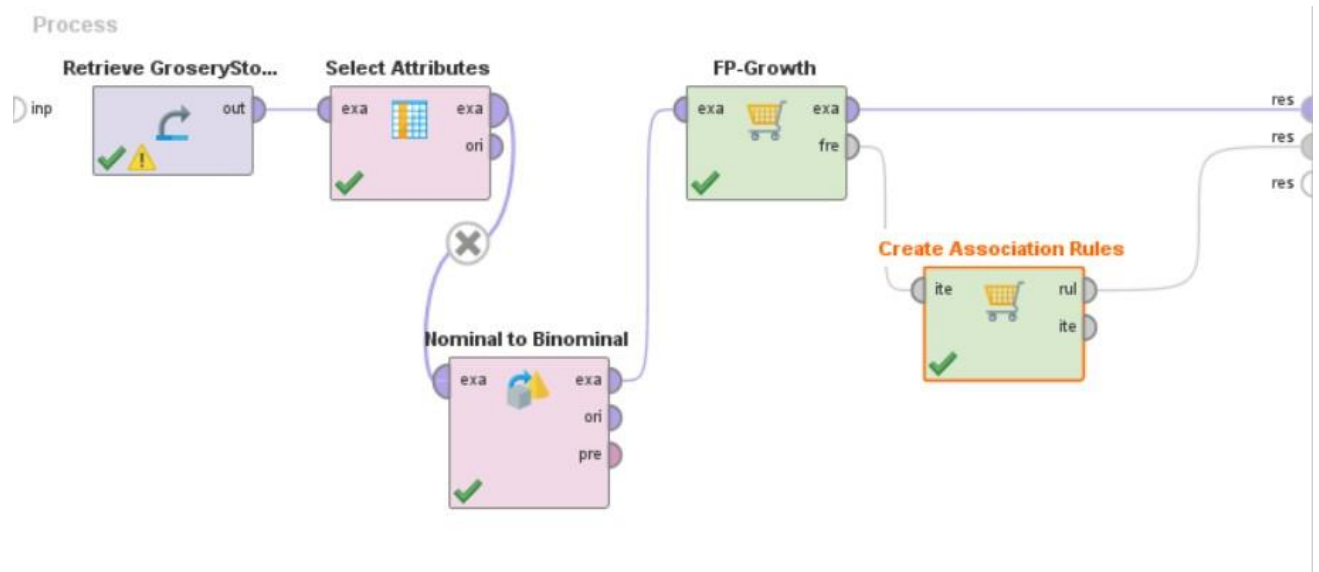
Model associations between products by determining sets of items frequently purchased together & building association rules to derive recommendations.

Conclusion:-

Thus, we learn that to find frequently occurring items from given data & generate strong association rules using support & confidence thresholds using a-priori algorithm.



## Output:



Result History

AssociationRules (Create Association Rules) X

### AssociationRules

Association Rules

- [BREAD] --> [MILK] (confidence: 0.857)
- [TEA] --> [MILK] (confidence: 0.857)
- [COFFEE] --> [MILK] (confidence: 0.875)
- [SUGER] --> [MILK] (confidence: 1.000)
- [BOURNVITA] --> [MILK] (confidence: 1.000)
- [BOURNVITA] --> [BISCUIT] (confidence: 1.000)
- [BISCUIT, SUGER] --> [MILK] (confidence: 1.000)
- [BOURNVITA] --> [MILK, BISCUIT] (confidence: 1.000)
- [MILK, BOURNVITA] --> [BISCUIT] (confidence: 1.000)
- [BISCUIT, BOURNVITA] --> [MILK] (confidence: 1.000)
- [COFFEE, SUGER] --> [MILK] (confidence: 1.000)
- [MILK, CORNFLAKES] --> [BREAD] (confidence: 1.000)
- [MILK, MAGGI] --> [TEA] (confidence: 1.000)
- [TEA, MAGGI] --> [MILK] (confidence: 1.000)
- [COFFEE, SUGER] --> [BISCUIT] (confidence: 1.000)
- [COFFEE, CORNFLAKES] --> [BREAD] (confidence: 1.000)
- [COFFEE, SUGER] --> [MILK, BISCUIT] (confidence: 1.000)
- [MILK, COFFEE, SUGER] --> [BISCUIT] (confidence: 1.000)
- [BISCUIT, COFFEE, SUGER] --> [MILK] (confidence: 1.000)

Activate Windows