Association Rules

Bert Gollnick

Association Rules

Introduction

- Goal is
  - to find associations between different objects
  - to find patterns in objects
  - to derive strategies (store layout, catalogue design, customizing emails, recommender engines, …)
- Applicable for
  - Marketing
  - Clustering
  - Classification
  - Retailing

Association Rules

Principle

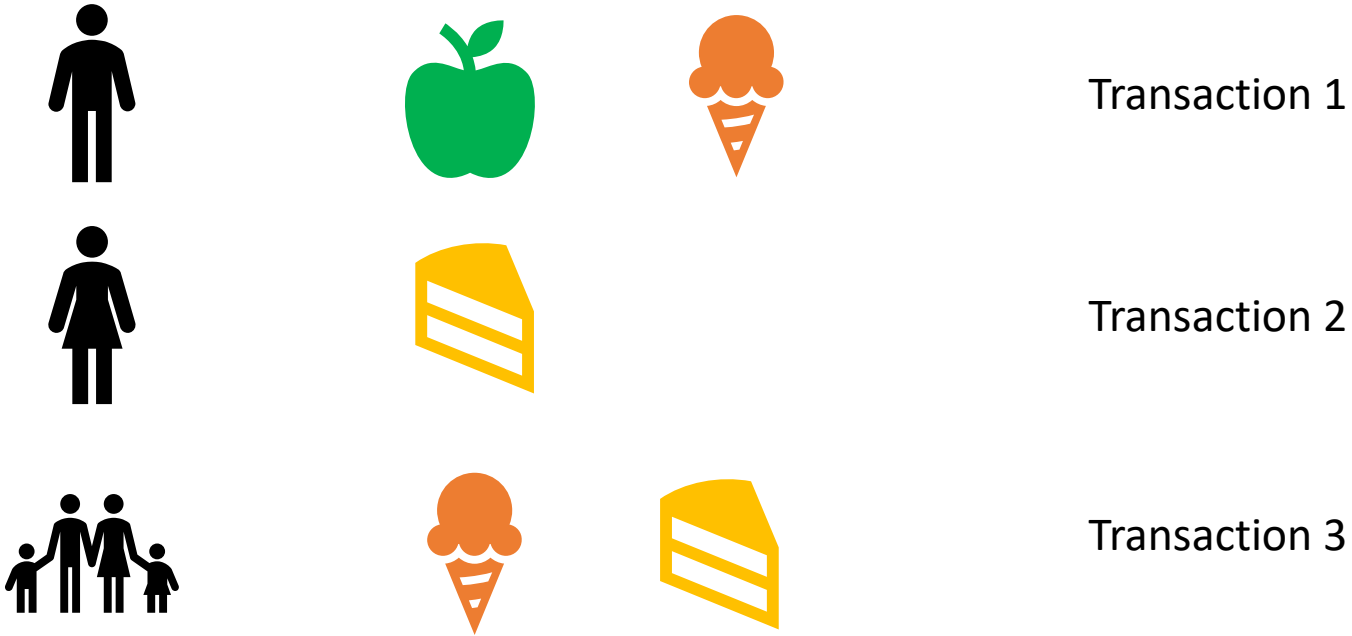
- Example: Supermarket

Product Portfolio → Item Set

Association Rules

Principle

- Input: Transactions



Association Rules

Principle

- Transactions

Transaction	Items
1	{apple, ice}
2	{cake}
3	{ice, cake}
…	{…}
N	{…}

Association Rules

Principle

- Output: Rules

[LHS]	[RHS]
{ice}	{cake}
{apple, cake}	{ice}
…	

Association Rules

Key Parameters: Support

- $Support(X) = \frac{Number\ Transactions\ including\ X}{Total\ Transactions\ Number} = \frac{count(X)}{N}$
- $Support(\{X\} \rightarrow \{Y\}) = \frac{Transactions\ containing\ X\ AND\ Y}{Total\ Transaction\ Number}$

- Support threshold can be defined to exclude infrequent transactions

Association Rules

Key Parameters: Support - Example

- Calculate Support(Apple → Pie)!
- Minimum Support = 0.4

- Count(apple&pie) = 3
- Total Number= 5
- Min Support excludes infrequent itemsets from further analysis

Transact. ID	
1	
2	 
3	 
4	  
5	  

Association Rules

Key Parameters: Confidence

- $Confidence(X \rightarrow Y) = \frac{Support(X,Y)}{Support(X)}$
- How often is Y bought, after X was bought?
- Problem: only considers popularity of X (not of Y)
  - If X and Y are popular, buying X and Y together is likely. So high confidence occurs by chance.

Association Rules

Key Parameters: Lift

- $Lift(X \rightarrow Y) = \frac{Support(X,Y)}{Support(X) * Support(Y)}$

- How often is Y bought, after X was bought?
- This takes popularity of Y into account.