Outline

- **Briefly explain association analysis**
- **How to implement association analysis in Weka**

Association Analysis

- **Association analysis (AA)** discovers the probability of the co-occurrence of items in a collection.
- **Association rules**: the relationships between co-occurring items.

Market-basket analysis

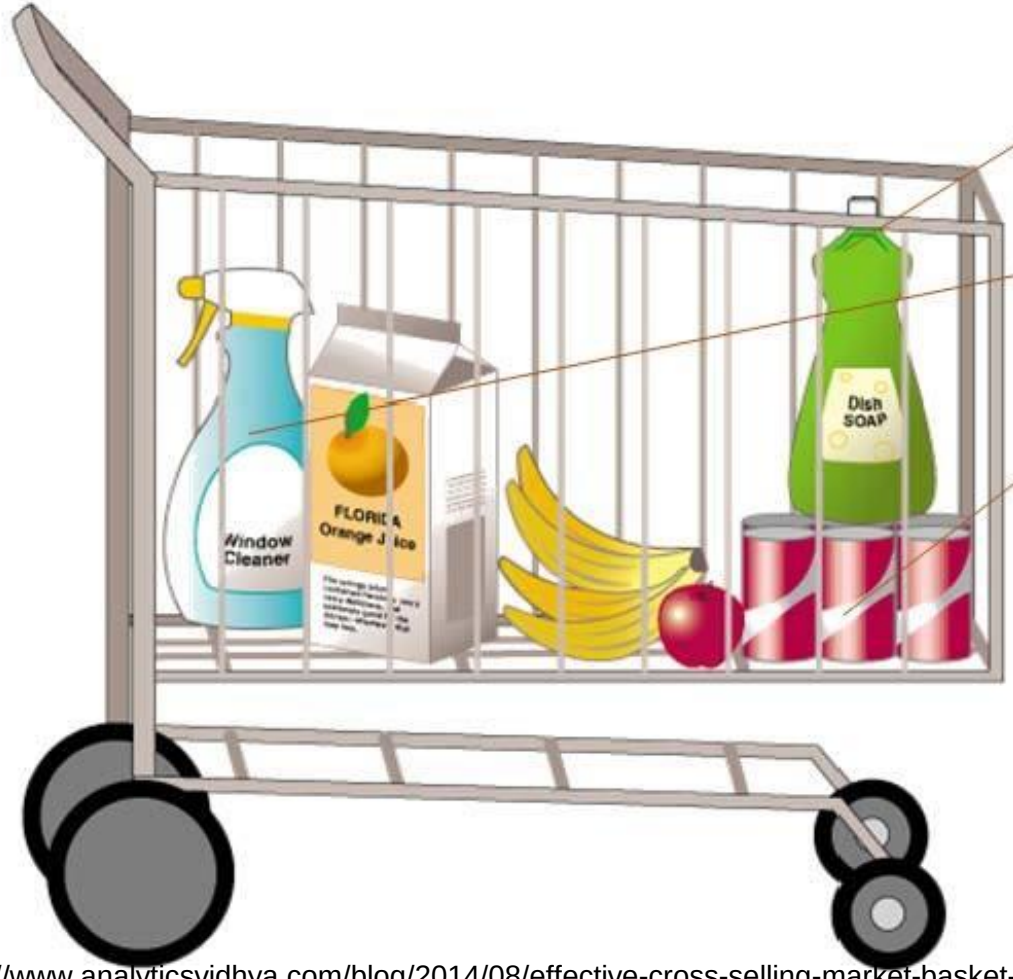
Valuable for direct marketing, sales promotions, and for discovering business trends. Market-basket analysis can also be used effectively for store layout, catalog design, and cross-sell.

Example: An association model might find that a user who bought products A and B is 70% likely to buy product C in the same session.

https://docs.oracle.com/cd/E18283_01/datamine.112/e16808/market_basket.htm

Market Basket Example

Example II



? Where should detergents be placed in the Store to maximize their sales?

? Are window cleaning products purchased when detergents and orange juice are bought together?

? Is soda typically purchased with bananas? Does the brand of soda make a difference?

? How are the demographics of the neighborhood affecting what customers are buying?

Image source: deepclimate.org

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

Association rules

$$Support = \frac{frq(X,Y)}{N}$$

$$Rule: X \Rightarrow Y \quad Confidence = \frac{frq(X,Y)}{frq(X)}$$

$$Lift = \frac{Support}{Supp(X) \times Supp(Y)}$$

Rule	Support	Confidence	Lift
$A \Rightarrow D$	2/5	2/3	10/9
$C \Rightarrow A$	2/5	2/4	5/6
$A \Rightarrow C$	2/5	2/3	5/6
$B \& C \Rightarrow D$	1/5	1/3	5/9

An example of Association Rules

1. Assume there are 100 customers.
2. 10 of them bought milk, 8 bought butter and 6 bought both of them.
3. bought milk => bought butter.
4. support = P(Milk & Butter) = 6/100 = 0.06.
5. confidence = support/P(Butter) = 0.06/0.08 = 0.75.
6. lift = confidence/P(Milk) = 0.75/0.10 = 7.5.

Please note the rule A ⇒ D differs from the rule D ⇒ A

https://ghumare64.medium.com/market-basket-analysis-using-association-rule-mining-2f87e4064ee7
https://www.saedsayad.com/association_rules.htm

Please pay attention to the data format requirement

Order	Product
1	Product 1
1	Product 2
1	Product 3
2	Product 2
2	Product 3
3	Product 2
3	Product 3
3	Product 4

Unacceptable

Product 1	Product 2	Product 3	…	Product n
1	1		..	1
	1		..	
	1	1	..	
		1	..	1
	1	1	..	1
			..	1

Acceptable