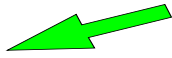


Frequent Pattern Mining & Association rule Mining

- What is frequent pattern analysis ? 
- Importance of frequent pattern analysis / mining
- Frequent patterns and association rules
- Frequent itemset mining methods

What Is Frequent Pattern Analysis?

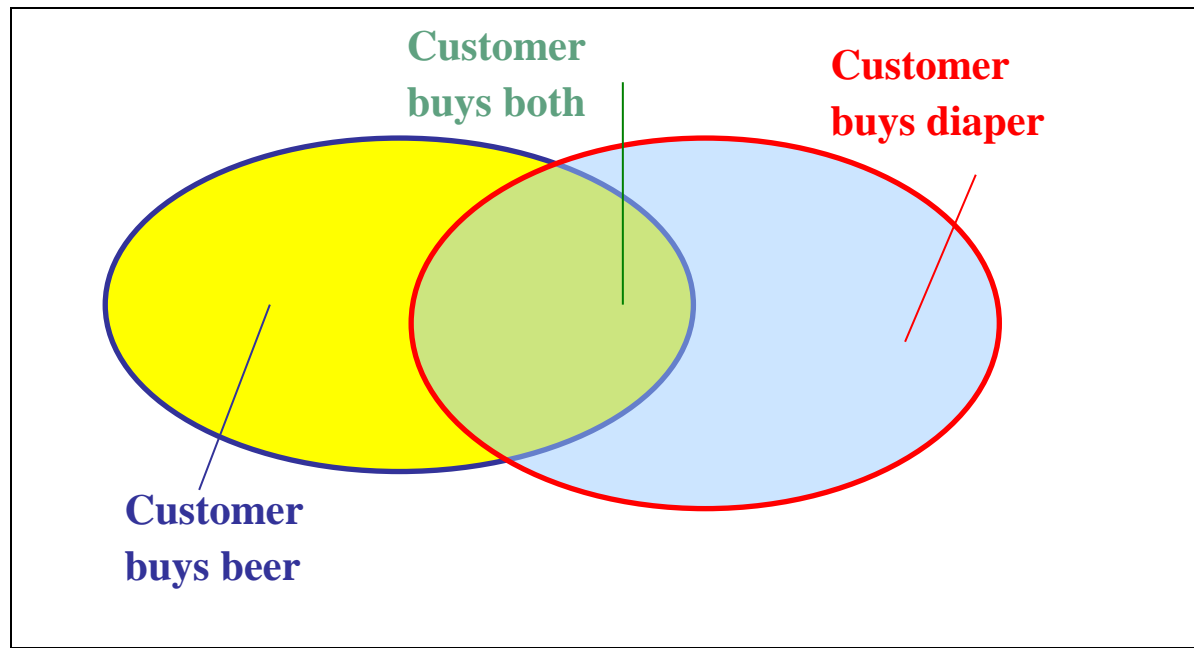
- **Frequent pattern:** a pattern (a set of items, subsequences etc.) that occurs frequently in a data set.
- Motivation: Finding regularities in data
 - What products were often purchased together?
 - What are the subsequent purchases after buying a specific Product?
- Applications: Market Basket analysis, catalog design, sale campaign analysis etc.

Importance of Frequent Pattern Mining

- Discloses an intrinsic and important property of data sets
- Forms the foundation for many essential data mining tasks
 - Association, correlation etc.
 - Sequential, structural (e.g., sub-graph) patterns
 - Pattern analysis in multimedia, time-series, and stream data
 - Classification: associative classification
 - Cluster analysis: frequent pattern-based clustering

Frequent Patterns and Association Rules

Transaction-id	Items bought
10	A, B, D
20	A, C, D
30	A, D, E
40	B, E, F
50	B, C, D, E, F



Frequent Patterns and Association Rules

- Itemset $X = \{x_1, \dots, x_k\}$
- Find all the rules $X \rightarrow Y$ with minimum support and confidence
 - **support**, s , **probability** that a transaction contains $X \cup Y$
 - **confidence**, c , **conditional probability** that a transaction having X also contains Y

Let $sup_{min} = 50\%$, $conf_{min} = 50\%$

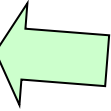
Freq. Pat.: $\{A:3, B:3, D:4, E:3, AD:3\}$

Association rules:

1. $A \rightarrow D$ (60%, 100%)
2. $D \rightarrow A$ (60%, 75%)

Scalable Frequent Itemset Mining Methods

- Apriori: A Candidate Generation-and-Test Approach
- Improving the Efficiency of Apriori
- FP Growth: A Frequent Pattern-Growth Approach



The Downward Closure Property and Scalable Mining Methods

- The **downward closure** property of frequent patterns
 - Any subset of a frequent itemset must be frequent
 - If **{beer, diaper, nuts}** is frequent, so is **{beer, diaper}**
 - i.e., every transaction having {beer, diaper, nuts} also contains {beer, diaper}
- Scalable mining methods: Two major approaches
 - Apriori Algorithm
 - Freq. pattern growth (FP growth)