PROJECT – MARKET BASKET INSIGHTS

Market Basket Analysis :

Market basket analysis is a strategic data mining technique used by retailers to enhance sales by gaining a deeper understanding of customer purchasing patterns. This method entails the examination of substantial datasets, such as historical purchase records, in order to unveil inherent product groupings and identify items that tend to be bought together.

By recognizing these patterns of co-occurrence, retailers can make informed decisions to optimize inventory management, devise effective marketing strategies, employ cross-selling tactics, and even refine store layout for improved customer engagement.

For example,

Market Basket Analysis Work

1. Collect data on customer transactions, such as the items purchased in each transaction, the time and date of the transaction, and any other relevant information.
2. Clean and preprocess the data, removing any irrelevant information, handling missing values, and converting the data into a suitable format for analysis.
3. Use association rules mining algorithms such as Apriori or FP-Growth to identify frequent item sets, sets of items often appearing together in a transaction.
4. Calculate the support and confidence for each frequent itemset, which expresses the likelihood of one item being purchased given the purchase of another item.
5. Generate association rules based on the frequent itemsets and their corresponding support and confidence values. Association rules express the likelihood of one item being purchased given the purchase of another item.
6. Interpret the results of the market basket analysis, identifying which items are frequently purchased together, the strength of the association between items, and any other relevant insights into customer behavior and preferences.
7. Use the insights from the market basket analysis to inform business decisions such as product recommendations, store layout optimization, and targeted marketing campaigns.

3 Types of Market Basket Analysis

1. It involves identifying frequent item sets and generating association rules that express the likelihood of one item being purchased with the purchase of another item. It is used to identify the relationships or associations between items in a transactional dataset.
2. This type of market basket analysis focuses on the order in which items are purchased in a transaction. It identifies frequent item sequences and generates sequential association rules describing the likelihood of one item sequence being followed by another.
3. This type of market basket analysis involves grouping similar items or transactions into clusters or segments based on their attributes. It helps to identify customer segments with similar purchasing behaviors, which can inform product recommendations and marketing strategies.

Applications of Market Basket Analysis

Industry	Applications of Market Basket Analysis
Retail	Identify frequently purchased product combinations and create promotions or cross-selling strategies
E-commerce	Suggest complementary products to customers and improve the customer experience
Hospitality	Identify which menu items are often ordered together and create meal packages or menu recommendations
Healthcare	Understand which medications are often prescribed together and identify patterns in patient behavior or treatment outcomes
Banking/Finance	Identify which products or services are frequently used together by customers and create targeted marketing campaigns or bundle deals
Telecommunications	Understand which products or services are often purchased together and create bundled service packages that increase revenue and improve the customer experience

Association Rule for Market Basket Analysis

Let I = {I1, I2,…, Im} be an itemset. These itemsets are called antecedents. Let D, the data, be a set of database transactions where each transaction T is a nonempty itemset such that **T ⊆ I**. Each transaction is associated with an identifier called a TID(or Tid). Let A be a set of items(itemset). T is the Transaction that is said to contain A if **A ⊆ T**. An **Association Rule** is an implication of form **A ⇒ B**, where **A ⊂ I, B ⊂ I**, and **A ∩B = φ**.

The rule **A ⇒ B** holds in the data set(transactions) D with supports, where 's' is the percentage of transactions in D that contain **A ∪ B** (i.e., the union of set A and set B, or both A and B). This is taken as the probability, **P(A ∪ B)**. Rule **A ⇒ B** has confidence **c** in the transaction set D, where c is the percentage of transactions in D containing **A** that also contains **B**. This is taken to be the conditional probability, like P(B|A). That is,

- ***support(A⇒ B) =P(A ∪ B)***
- ***confidence(A⇒ B) =P(B|A)***

Rules that satisfy both a minimum support threshold (called min sup) and a minimum confidence threshold (called min conf) are called "***Strong"***.

- ***Confidence(A⇒ B) = P(B|A) =***
- ***support(A ∪B) /support(A) =***
- ***support count(A ∪B) / support count(A)***

Generally, Association Rule Mining can be viewed in a two-step process:

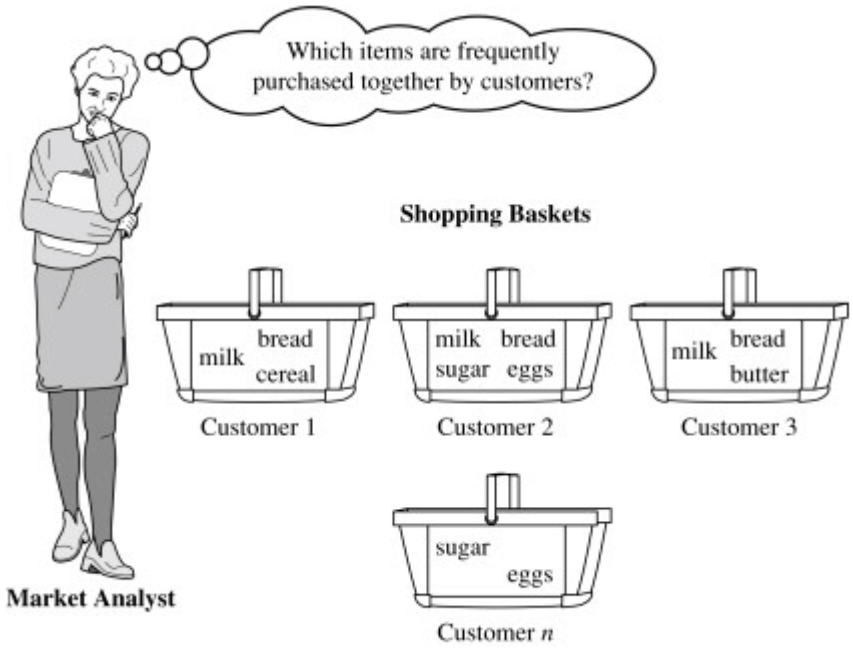
1. Find all Frequent itemsets: *By definition, each of these itemsets will occur at least as frequently as a pre-established minimum support count, min sup*.

2. Generate Association Rules from the Frequent itemsets: *By definition, these*
   rules must satisfy minimum support and minimum confidence.

Association Rule Mining

Association Rule Mining is primarily used when you want to identify an association between different items in a set and then find frequent patterns in a transactional database or relational database.

The best example of the association is as you can see in the following image.



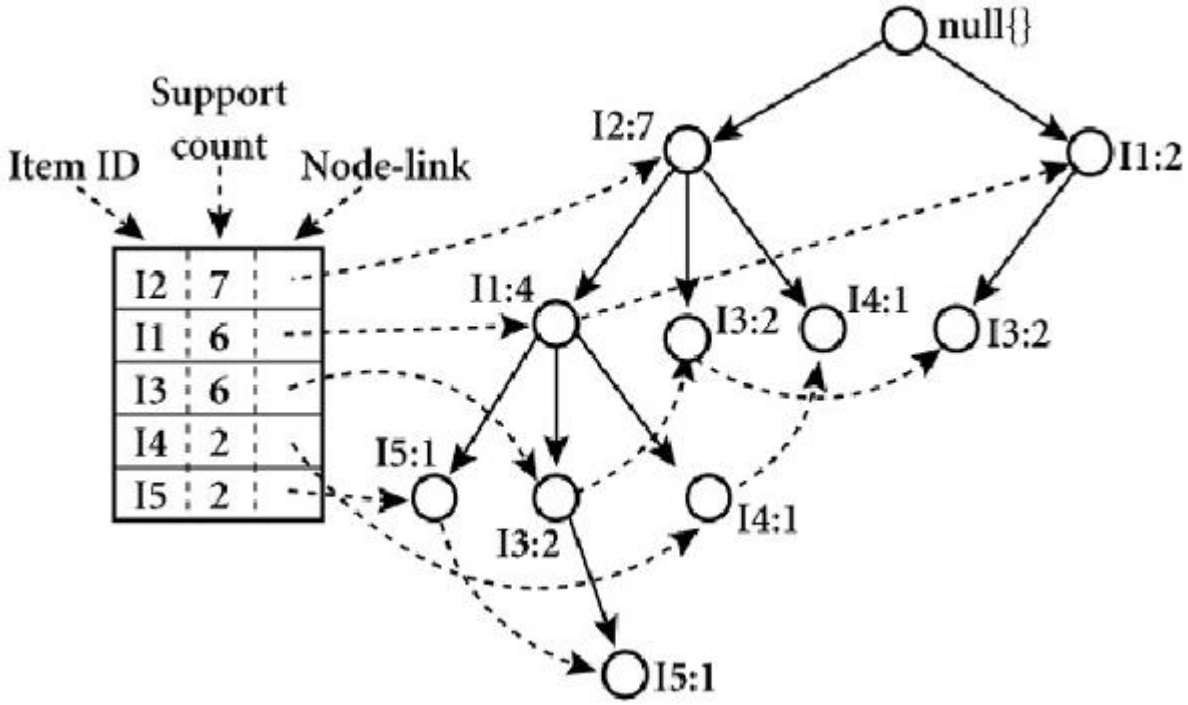
Algorithms Used in Market Basket Analysis

Apriori Algorithm

Apriori Algorithm is a widely-used and well-known Association Rule algorithm and is a popular algorithm used in market basket analysis. It is also considered accurate and overtop AIS and SETM algorithms. It helps to find

frequent itemsets in transactions and identifies association rules between these items. The limitation of the Apriori Algorithm is *frequent itemset generation*. It needs to scan the database many times, leading to increased time and reduced performance as a computationally costly step because of a large dataset. It uses the concepts of Confidence and Support.

For example,



Advantages of Market Basket Analysis

There are many advantages to implementing Market Basket Analysis in marketing. Market Basket Analysis (MBA) can be applied to data of customers from the point of sale (**PoS**) systems.

It helps retailers in the following ways:

- Increases customer engagement

- Boosts sales and increases RoI
- Improves customer experience
- Optimizes marketing strategies and campaigns
- Helps in demographic data analysis
- Identifies customer behavior and pattern

Market Basket Analysis From the Customers' Perspective

Let us take an example of market basket analysis from Amazon, the world's largest eCommerce platform. From a customer's perspective, Market Basket Analysis is like ***shopping at a supermarket***. Generally, it observes all items bought by customers together in a single purchase. Then it shows the most related products together that customers will tend to buy in one purchase.

Implementing Market Basket Analysis in Python

The Method

Here are the steps involved in using the apriori algorithm to implement MBA:

1. First, define the minimum support and confidence for the association rule.

2. Find out all the subsets in the transactions with higher support(sup) than the minimum support.
3. Find all the rules for these subsets with higher confidence than minimum confidence.
4. Sort these association rules in decreasing order.
5. Analyze the rules along with their confidence and support.

Conclusion

we discussed Market Basket Analysis and learned the steps to implement it from scratch using Python. We then implemented Market Basket Analysis using Apriori Algorithm. We also looked into the various uses and advantages of this algorithm in Market Basket Analysis.