Market basket analysis with Apriori algorithm

Introduction

The retailer wants to target customers with suggestions on itemset that a customer is most likely to purchase .I was given dataset contains data of a retailer; the transaction data provides data around all the transactions that have happened over a period of time. Retailer will use result to grove in his industry and provide for customer suggestions on itemset, we be able increase customer engagement and improve customer experience and identify customer behaviour. I will solve this problem with use Association Rules type of unsupervised learning technique that checks for the dependency of one data item on another data item.

Problem Statement:

The problem is to perform market basket analysis on a provided dataset to unveil hidden patterns and associations between products. The goal is to understand customer purchasing behavior and identify potential cross-selling opportunities for a retail business. This project involves using association analysis techniques, such as Apriori algorithm, to find frequently co-occurring products and generate insights for business optimization.

Understanding the Problem:

Market basket analysis is a crucial aspect of retail strategy. By understanding which products are frequently bought together, businesses can optimize their inventory, marketing strategies, and store layouts. The Apriori algorithm, in particular, is adept at handling large datasets and discovering these patterns efficiently.

Design Thinking:

1. Data Understanding:

The first step involves understanding the dataset provided. This includes grasping the structure of the data, the variables available, and the nature of the transactions. This understanding is fundamental for subsequent preprocessing and analysis.

2. Data Preprocessing:

Data preprocessing includes handling missing values, removing duplicates, and transforming the data into transactional format where each transaction comprises a list of purchased items.

3.Applying the Apriori Algorithm:

Apriori algorithm is used to perform association analysis on the preprocessed data. Appropriate thresholds are set to filter out significant associations. This step will generate frequent itemsets which are groups of items frequently bought together and association rules which represent relationships between products.

4. Interpretation of Results:

The generated association rules are interpreted to understand the relationships between products. The products that are commonly bought together and strength of those associations are identified. This interpretation forms the basis for deriving insights into customer behavior and preferences.

5. Visualization:

Visualization tools like heatmaps, network graphs, or simple bar charts can be employed to represent the discovered patterns and associations effectively.

6. Business Recommendations:

Based on the interpreted results and visualizations actionable recommendations can be provided to the retail business. These recommendations can include optimizing product placements in stores, creating bundled offers, or designing targeted marketing campaigns.

About Association Rule

Association Rule is most used when you are planning to build association in different objects in a set. It works when you are planning to find frequent patterns in a transaction database. It can tell you what items do customers frequently buy together and it allows retailer to identify relationships between the items.