INTRODUCTION

Frequent pattern mining finds the most frequent and relevant patterns in large dataset. Frequent patterns are patterns that appear frequently in a dataset. Finding frequent patterns plays an essential role in mining associations, correlations, causal structures, and many other interesting relationships among data. These interesting relationships mined used to make various decisions in the present and plan for the future. Thus, frequent pattern mining is an important data mining task. Association rule mining is a popular method for discovering interesting co-occurring relations between variables in large dataset.

For example, from the sales data of stationery frequent pattern mining finds pencil, paper and rubber as frequent pattern. From this frequent pattern association rule mining find interesting association of the form- customers that buy a pencil and paper are likely to buy a rubber. Such association can be used as the basis for decisions about marketing activities such as, to recommend next item to the customer, promotional pricing, product placements, catalog design, cross-marketing, sale campaign analysis, etc. In addition to the above example from market basket analysis association rule mining are employed today in many application areas including web usage mining, intrusion detection, software bugs, and bioinformatics.

This chapter introduces the basic concepts of frequent patterns, and association rule mining and describes matrices to evaluate the interestingness of the mined association rules such as support, confidence, and correlation analysis (e.g., lift). This chapter mainly focuses on various frequent pattern mining algorithms such as Apriori, FP-growth algorithm, and association rule mining from the frequent patterns. At the end of this chapter, we study the

FREQUENT PATTERNS

Frequent patterns are itemsets, subsequences, or substructures that appear in a data set with frequency no less than a user-specified threshold. For example, a set of items, such as Paper and Pencil that appear frequently together in a stationery transaction data set is a frequent itemset. A subsequence, such as buying first a PC, then a digital camera, and then a memory card, if it occurs frequently in a shopping history database, is a frequent sequential pattern. A substructure can refer to different structural forms, such as subgraphs, subtrees, or sublattices, which may be combined with itemsets or subsequences. If a substructure occurs frequently in a graph dataset, it is called a frequent structural pattern. Finding frequent patterns plays an essential role in mining associations, correlations, causal structures, and many other interesting relationships among data. Moreover, it helps in data classification, clustering, and other data mining tasks as well.

MARKET BASKET ANALYSIS

Market basket analysis is an example of a frequent pattern mining mainly used by retailers to understand customer purchase behaviors. In market basket analysis retailers determines what items are frequently bought together or placed in the same basket by customers. Based on the combinations of products that frequently co-occur in transactions retailer uncovers associations between products. These uncovered associations are then used by retailers to make purchase suggestions to consumers. For example, when a person buys a particular model of smartphone, the

retailer may suggest other products such as phone cases, screen protectors, memory cards or other accessories for that particular phone. This is due to the frequency with which other consumers bought these items in the same transaction as the phone. Similarly, stationery retailers also, make his recommendation of next item to the customer based on the baskets similar to shown in figure 5.1 below.

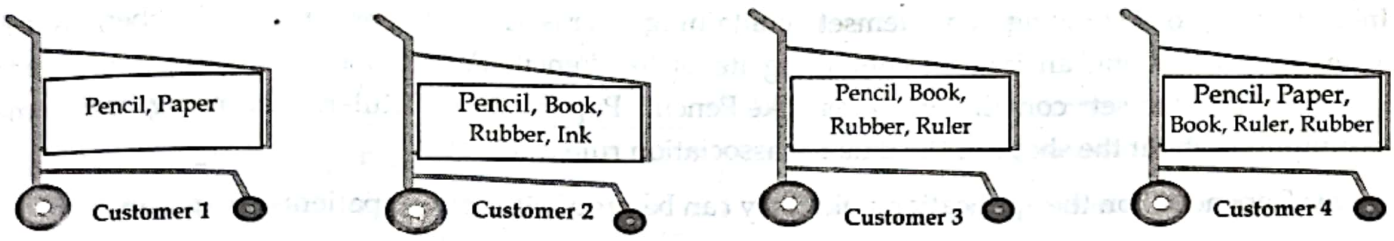

Figure 5.1: Market baskets for Stationary

Market basket analysis only uses transactions with more than one item, as no associations can be made with single purchases. Besides this it does not consider the order of items either within a transaction or across transaction

Benefits of Market Basket Analysis

Market basket analysis can increase sales and customer satisfaction. Using data to determine what products are often purchased together, retailers can optimize product placement, offer special deals and create new product bundles to encourage further sales of these combinations. These improvements can generate additional sales for the retailer, while making the shopping experience more productive and valuable for customers. By using market basket analysis, customers may feel a stronger sentiment or brand loyalty toward the company.

Market basket analysis isn't limited to shopping baskets. Other areas where the technique is used include e-commerce web sites, analysis of credit card purchases, analysis of telephone calling patterns, Identification of fraudulent insurance claims, analysis of telecom service purchases etc.

FREQUENT ITEMSETS

Transaction Dataset (D): In association rule mining, we are usually interested in the absolute number of customer transactions, also called baskets that contain a particular set of items, usually products. A typical application of association rule mining is the analysis of consumer buying behavior in supermarkets where they record the contents of shopping carts brought to register for checkout. These transaction data are normally consisting of tuples of the form as shown in below table 5.1.

Table 5.1: Market Basket Transactions

TID	Items
1	Pencil, Paper
2	Pencil, Book, Rubber, Ink
3	Paper, Book, Rubber, Ruler
4	Pencil, Paper, Book, Rubber
5	Pencil, Paper, Book, Ruler