**Introduction to Data Science**

**Week-6**

**Long Descriptive Questions**

**1, Explain Association Rule Mining with Example**

Association Rule Mining is a data mining technique used to discover interesting relationships, patterns, or associations within large datasets. It is commonly applied in various domains, such as market basket analysis, recommendation systems, and fraud detection. Association rules reveal how items or variables tend to co-occur in transactions or datasets.

The most well-known application of association rule mining is market basket analysis, where it helps identify patterns in customer purchasing behavior. Let's explain the concept with an example:

Example Market Basket Analysis

Consider a dataset from a grocery store where each transaction represents items purchased by a customer. The dataset might look like this:

Transaction 1- Bread, Milk, Eggs

Transaction 2- Bread, Diapers, Beer, Milk

Transaction 3- Milk, Diapers, Beer, Coke

Transaction 4- Bread, Milk, Diapers, Beer

Transaction 5- Bread, Milk, Diapers, Coke

Now, let's perform association rule mining to identify interesting associations or patterns in customer purchases.

Support: Support is a measure of how frequently an itemset (a combination of items) appears in the dataset. It is calculated as the number of transactions containing the itemset divided by the total number of transactions. Support helps identify frequently occurring itemsets.

Confidence: Confidence measures the likelihood that when one item (A) is purchased, another item (B) will also be purchased. It is calculated as the support for the itemset {A, B} divided by the support for itemset A. Confidence helps identify strong associations between items.

Example Rules

1. **{Milk} -> {Bread}**

Support: 3/5 (3 transactions with Milk and Bread)

Confidence: 3/4 (3 out of 4 transactions with Milk also have Bread)

1. **{Diapers} -> {Beer}**

Support: 2/5 (2 transactions with Diapers and Beer)

Confidence: 2/3 (2 out of 3 transactions with Diapers also have Beer)

1. **{Bread, Milk} -> {Diapers}**

Support: 2/5 (2 transactions with Bread, Milk, and Diapers)

Confidence: 2/3 (2 out of 3 transactions with Bread and Milk also have Diapers)

Based on these rules, we can make several observations:

Rule 1 suggests that customers who buy Milk are likely to buy Bread as well (high confidence).

Rule 2 indicates that customers who buy Diapers are likely to buy Beer (high confidence).

Rule 3 shows that customers who buy both Bread and Milk are likely to buy Diapers (high confidence).

These association rules can be valuable for marketing and inventory management. For example, the store can use this information to plan product placements, promotions, and product bundling strategies to increase sales and customer satisfaction.