

Introduction to Apriori Algorithm

Understanding Market Basket Analysis and Association Rule Mining

Introduction

- ▶ The **Apriori Algorithm** is a popular data mining algorithm used for **association rule learning**.
- ▶ It is widely used in market basket analysis to identify frequent item sets and derive strong association rules.
- ▶ The algorithm is based on the **Apriori property**, which states that a subset of a frequent itemset must also be frequent.

Advantages of Apriori Algorithm

- ▶ Simple and easy to implement.
- ▶ Effective for small datasets.
- ▶ Useful for discovering interesting patterns.
- ▶ Supports Boolean association rules.

Disadvantages of Apriori Algorithm

- ▶ Computationally expensive.
- ▶ High memory usage.
- ▶ Redundant rule generation.
- ▶ Not efficient for continuous data.

Applications of Apriori Algorithm

- ▶ Market Basket Analysis.
- ▶ Medical Diagnosis.
- ▶ Fraud Detection.
- ▶ Recommendation Systems.
- ▶ Web Usage Mining.

Definitions – Support Confidence

Support: Measures how frequently an itemset appears in the dataset.

$$Support(A) = \frac{\text{Frequency of transactions containing } A}{\text{Total transactions}} \quad (1)$$

Confidence: Measures how often items in B appear in transactions that contain A.

$$Confidence(A \Rightarrow B) = \frac{Support(A \cup B)}{Support(A)} \quad (2)$$

Conclusion

- ▶ The **Apriori Algorithm** is a fundamental technique for **association rule mining**.
- ▶ It is useful for discovering patterns but has computational limitations.
- ▶ Widely applied in industries like retail, healthcare, and finance.