

MSBA 511: Data Mining for Business Analytics

Conceptual Homework 1: Association Rules

Objective

Be able to:

- Apply the steps of the Apriori algorithm to find frequent itemsets and valid association rules based on defined thresholds.

Overview

In this assignment, you will analyze transaction data to generate association rules using the Apriori algorithm. Your goal is to identify patterns in the data with a **minimum support of 30%** and a **minimum confidence of 50%**.

Data

The dataset includes transaction IDs and items purchased together:

tran_id	items
1	Bread, Butter, Soap
2	Milk, Butter
3	Peas, Milk, Soap
4	Bread, Peas, Butter, Soap
5	Bread, Peas, Milk, Soap
6	Bread, Butter, Soap
7	Peas, Butter, Soap
8	Bread, Peas, Milk
9	Bread, Milk, Butter, Soap
10	Bread, Peas, Soap

Tasks

1. For the minimum support threshold, identify all frequent itemsets.
2. Based on the identified frequent itemsets, identify all valid association rules based on the minimum confidence threshold.

Submission Requirements

Submit a PDF document to Gradescope that includes the following:

1. A table of frequent itemsets with support values.
2. A table of association rules with:
 - a. Antecedent
 - b. Consequent
 - c. Support
 - d. Confidence
 - e. Lift
3. Detailed analysis steps for each step. You may include screenshots or graphics in addition to a textual summary.
4. Interpret the results in a business context.
 - a. For each rule, explain what it means in the context of this data and highlight any interesting or actionable insights you observe.