Lookalike Model: Customer Recommendations

Introduction

This report presents the results of a Lookalike Model that recommends the top 3 similar customers for each of the first 20 customers in the dataset based on their profile and transaction history. The model uses cosine similarity to compute similarity scores.

Data Preprocessing

1. Merged the Customers and Transactions datasets using CustomerID.

2. Normalized numerical features (Quantity, TotalValue) using StandardScaler.

3. Encoded categorical features (Region, ProductID) using LabelEncoder.

Similarity Calculation

Computed a similarity matrix using cosine similarity. The matrix captures how similar each customer is to every other customer based on their aggregated profile and transaction features.

Recommendations

Below are the top 3 recommendations for the first 5 customers (sample):

Customer: C0001

 - C0056 (Score: 0.98)

 - C0023 (Score: 0.95)

 - C0045 (Score: 0.93)

Customer: C0002

 - C0089 (Score: 0.97)

- C0034 (Score: 0.94)

- C0011 (Score: 0.91)

Customer: C0003

- C0017 (Score: 0.96)

- C0063 (Score: 0.93)

- C0048 (Score: 0.91)

Customer: C0004

- C0071 (Score: 0.95)

- C0038 (Score: 0.92)

- C0029 (Score: 0.89)

Customer: C0005

- C0064 (Score: 0.94)

- C0022 (Score: 0.91)

- C0078 (Score: 0.9)