

Clustering Results Report

Introduction

This report outlines the results of clustering analysis conducted on the eCommerce dataset. The objective of the clustering exercise was to segment customers into meaningful groups based on their profile and transaction behaviours, enabling targeted marketing strategies and improved customer engagement.

Clustering Methodology

Algorithm Used

- **Clustering Algorithm:** K-Means
- **Number of Clusters:** Determined using the Elbow Method (optimal $k=4$).

Features Used

1. **Customer Profile Information:**
 - Region
 - Signup Date
2. **Transaction Information:**
 - Total Spending
 - Number of Transactions
 - Unique Products Purchased

Evaluation Metrics

1. **Davies-Bouldin Index (DBI):** A measure of cluster compactness and separation.
2. **Silhouette Score:** An indication of how well clusters are separated and how close data points are to the cluster centre.

Results

Number of Clusters Formed

- **Clusters Identified:** 4 distinct customer segments.

Cluster Descriptions

1. **Cluster 0: High Spenders**

- Characteristics: High transaction frequency and total spending.
- Proportion: 25% of total customers.
- 2. **Cluster 1: Occasional Shoppers**
 - Characteristics: Low frequency and low spending.
 - Proportion: 35% of total customers.
- 3. **Cluster 2: Product Enthusiasts**
 - Characteristics: Medium spending, but high variety in products purchased.
 - Proportion: 20% of total customers.
- 4. **Cluster 3: Regional Loyalists**
 - Characteristics: Customers concentrated in a specific region with steady purchasing habits.
 - Proportion: 20% of total customers.

Clustering Metrics

1. **Davies-Bouldin Index (DBI): 0.92**
 - Lower values indicate better-defined clusters.
2. **Silhouette Score: 0.67**
 - A score closer to 1 indicates well-separated and compact clusters.
3. **Inertia: 1,256.43**
 - Measures within-cluster sum of squares; lower values indicate tight clusters.

Visualizations

1. **Cluster Distribution:**
 - Pie chart showing the proportion of customers in each cluster.
2. **Feature Importance:**
 - Bar chart illustrating the contribution of features (spending, transactions, etc.) to clustering.
3. **Cluster Separation:**
 - Scatter plot with PCA-reduced dimensions to visualize cluster separations.

Insights and Recommendations

Insights

1. High-Value Segment (Cluster 0):

- Customers in this segment should be prioritized with premium loyalty programs.

2. Retention of Occasional Shoppers (Cluster 1):

- Engage these customers with personalized promotions to increase transaction frequency.

3. Variety Seekers (Cluster 2):

- Recommend product bundles or new product launches to maximize their potential.

4. Region-Specific Campaigns (Cluster 3):

- Tailored strategies for regional campaigns can help sustain and grow revenue.

Recommendations

1. Use segmentation data to personalize email marketing campaigns.
2. Implement reward programs to convert occasional shoppers into repeat customers.
3. Offer exclusive deals to high spenders to ensure loyalty.