

Summary of Model Implementation

- 1. Feature Engineering:** We integrate customer profiles, transaction behaviors, and product preferences into a unified feature set. All features are standardized and converted to numerical values to ensure they are suitable for clustering.
- 2. Optimal Cluster Selection:** The Davies-Bouldin Index is used to identify the optimal number of clusters. Additionally, the Silhouette Score is employed to validate the results and increase confidence in our findings.
- 3. Clustering:** K-means clustering is applied to group customers into segments based on their feature vectors.
- 4. Visualization:** Principal Component Analysis (PCA) is utilized to reduce the dimensionality of the data, allowing us to visualize the clusters in a two-dimensional space.
- 5. Analysis and Reporting:** We analyze the characteristics of each cluster and compile the results for business use. This implementation ensures that customer segmentation is data-driven, interpretable, and actionable for strategic business decisions.

Clustering Visualization Image

