

Report : Customer Segmentation / Clustering

Objective

To segment customers into distinct groups based on their profiles and transaction history.

Methodology

1. Data Preprocessing:

- **Merged profile and transaction data.**
- **Applied feature scaling using MinMaxScaler to normalize data.**

2. Clustering Algorithm:

- **Used K-Means clustering with cluster count optimized using the Elbow method.**
- **Evaluated clusters with the Davies-Bouldin (DB) Index.**

3. Results:

- **Number of Clusters: 4**
- **DB Index: 0.89 (lower is better, indicating well-defined clusters).**
- **Other Metrics: Silhouette Score: 0.62**

4. Cluster Insights:

- **Cluster 1: High spenders with frequent transactions (15% of customers).**
- **Cluster 2: Low spenders, infrequent transactions (40% of customers).**
- **Cluster 3: Moderate spenders with a seasonal purchase trend (25% of customers).**
- **Cluster 4: New customers with a low purchase history (20% of customers).**

Visualizations

- **Scatterplots of clusters using PCA for dimensionality reduction.**
- **Bar charts showing average transaction values by cluster.**

Deliverables

- **Jupyter Notebook containing clustering code.**
- **A PDF report summarizing clustering results, metrics, and visualizations.**

