

Customer Segmentation Using Machine Learning

Introduction

This project focuses on customer segmentation using Machine Learning techniques to help businesses understand different customer groups based on behavior and demographics. Through Exploratory Data Analysis, clustering algorithms, and dimensionality reduction, this project offers actionable insights that support targeted marketing, personalization, and improved customer engagement.

1. Data Collection

A structured dataset of 20,000 customers was used containing Age, Income, Gender, Spending Score, Transactions, and Recency.

2. Data Cleaning

Removed duplicates, handled missing values, corrected data types, and filtered outliers.

3. Exploratory Data Analysis

Visualized distributions of key variables such as Age, Income, Spending Score, Gender Count, and relationships between features.

4. Feature Scaling

Applied StandardScaler to normalize values for optimal clustering performance.

5. Cluster Optimization

Used Elbow Method and Silhouette Score to determine best number of clusters (K).

6. K-Means Clustering

Performed clustering using K-Means and labeled each customer with a segment.

7. PCA Visualization

Used Principal Component Analysis to reduce dimensions and visualize clusters in 2D.

8. Cluster Profiling

Identified common behavior patterns in each segment such as High Spenders, Budget Customers, Premium Shoppers, etc.

9. Conclusion

This segmentation enables businesses to design personalized marketing strategies, improve customer retention, and increase revenue.