Abstract

This project explores customer segmentation through the application of unsupervised learning, specifically using the K-Means Clustering algorithm. By analyzing the Mall Customers Dataset, which includes features such as age, annual income, and spending score, our objective is to categorize customers into distinct groups based on their shopping behaviors. This segmentation provides insights for targeted marketing strategies, allowing businesses to understand different customer types and cater to their specific needs.

We implement the K-Means Clustering algorithm to identify optimal customer clusters, determined by the elbow method, which measures the point where adding more clusters no longer significantly reduces within-cluster variance. For enhanced interpretability, we use Principal Component Analysis (PCA) to visualize clusters in a two-dimensional space, making the differences between customer groups more apparent.

Through this model, we anticipate uncovering distinct customer segments, such as "High-Income High-Spenders" and "Low-Income Moderate-Spenders." These insights offer valuable guidance for personalized marketing efforts and strategic decision-making, highlighting the effectiveness of unsupervised learning in real-world business applications.