Customer Segmentation Using K-Means and PCA

Introduction

Customer segmentation is a key marketing strategy that helps businesses understand customer behavior and target them effectively. In this study, K-Means clustering was applied to the Mall Customer dataset using age, annual income, and spending score.

Clustering Approach

- **K-Means** was used to segment customers into distinct groups.

- The **Elbow Method** indicated that **5 clusters** is the optimal choice.

- **PCA (Principal Component Analysis)** was applied to reduce dimensionality and visualize clusters in 2D space.

Cluster Characteristics

The identified clusters show distinct customer behaviors:

- **High Income – High Spending:** Premium customers; ideal for luxury products.

- **High Income – Low Spending:** Potential customers; require targeted promotions.

- **Low Income – High Spending:** Value-seeking customers; respond well to discounts.

- **Low Income – Low Spending:** Budget-conscious customers.

- **Moderate Income – Moderate Spending:** Average customers; suitable for standard marketing.

Visualization Insights

- PCA visualization clearly separates customer groups.

- Income vs Spending Score plots show strong differentiation between high-value and low-value customers.

Business Marketing Strategy Impact

Customer segmentation enables:

- Personalized marketing campaigns

- Efficient allocation of marketing budgets

- Improved customer retention

- Better product recommendation strategies

Conclusion

The combination of K-Means clustering and PCA provides meaningful customer segments that can significantly enhance business decision-making. Understanding these clusters allows companies to tailor marketing strategies, improve customer satisfaction, and maximize revenue.