Customer Journey Analysis Using Clustering and Dimensionality Reduction

# 1. Introduction

Customer journey analysis is a critical process in understanding how customers interact with a business across various touchpoints. This analysis helps organizations enhance user experience, improve retention, and increase conversion rates. By applying clustering and dimensionality reduction techniques, we can uncover hidden patterns and segment customers based on their behavior.

# 2. Objective

The objective of this research is to perform customer journey analysis using unsupervised machine learning techniques—specifically, clustering (e.g., K-Means) and dimensionality reduction (e.g., PCA, t-SNE). The goal is to identify distinct customer segments and visualize high-dimensional journey data in a comprehensible manner.

# 3. Methodology

The methodology involves the following steps:  
1. Data Collection: Gathering customer journey data, including page visits, time spent, purchases, etc.  
2. Data Preprocessing: Cleaning and normalizing the data for analysis.  
3. Dimensionality Reduction: Applying techniques like PCA and t-SNE to reduce data dimensionality while preserving significant information.  
4. Clustering: Using K-Means or hierarchical clustering to identify customer segments.  
5. Visualization and Interpretation: Visualizing clusters and interpreting insights for business strategy.

# 4. Tools and Technologies

The following tools and technologies are used in the analysis:  
- Python (pandas, scikit-learn, matplotlib, seaborn)  
- Google Colab  
- Jupyter Notebooks  
- Dimensionality Reduction: PCA, t-SNE  
- Clustering Algorithms: K-Means, DBSCAN, Agglomerative Clustering

# 5. Results and Findings

The clustering analysis reveals several distinct groups of customers based on their journey behavior. Dimensionality reduction helps in visualizing these groups in 2D or 3D space. The insights from clusters can be used to target marketing campaigns, personalize user experiences, and optimize customer support.

# 6. Conclusion

Customer journey analysis using clustering and dimensionality reduction is an effective approach to understand user behavior. It aids businesses in making informed decisions, enhancing customer satisfaction, and increasing ROI. Future work may involve applying deep learning models or real-time analytics to further refine the insights.