CASE STUDY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

18CSC301J – MACHINE LEARNING

CASE STUDY

CO4 - CUSTOMER SEGMENTATION USING CLUSTERING TECHNIQUES

Student Register Number : 927622BCS031

Student Name : HARIHARAN S

III YEAR / V Semester /A Section

(2022-2026 BATCH)

Academic year: 2024-2025 (ODD Semester)

Submission Date:18/11/2024 Mark:

1. CUSTOMER SEGEMENTION USING CLUSTERING TECHNIQUES

Program:

Customer segmentation is a strategic approach where a business divides its customer base into distinct groups or segments based on shared characteristics. This enables personalized marketing, product recommendations, and improved customer service. Clustering techniques in machine learning are particularly useful for this because they allow automatic discovery of groups in data without predefined labels. Here's an example of how a business might use clustering for customer segmentation.

Objective

The objective of this case study is to segment customers based on purchasing behaviour, demographics, and engagement, enabling the business to:

1. Develop targeted marketing strategies.
2. Improve customer satisfaction and retention.
3. Personalize product recommendations and services.

Step-by-Step Process for Customer Segmentation

1. **Data Collection**
   - **Customer Data**: Gather customer information such as demographics (age, gender, income level), purchasing history, frequency of purchases, average order value, and engagement metrics (e.g., website visits, email interactions).
   - **Data Sources**: The data can be collected from CRM systems, transaction databases, and web analytics platforms.
2. **Data Preprocessing**
   - **Cleaning**: Remove duplicates, handle missing values, and address outliers.

- **Feature Engineering**: Create meaningful variables or features. For instance, calculate the average purchase amount or the frequency of interactions with the brand.
  .

3. **Choosing Clustering Algorithm**
   - There are several clustering algorithms commonly used in customer segmentation:
     - **K-Means Clustering**: This algorithm is widely used for its simplicity and effectiveness. It groups customers into kkk clusters based on the Euclidean distance between them.
     - **Hierarchical Clustering**: Builds a hierarchy of clusters and can be useful when the number of clusters is not predetermined.
     - **DBSCAN (Density-Based Spatial Clustering of Applications with Noise)**: Useful when clusters are of varying shapes and sizes, as it groups based on density.
4. **Selecting Features for Clustering**
   - Typical features for clustering might include:
     - **Recency**: Time since the last purchase.
     - **Frequency**: Number of purchases over a specified period.
     - **Monetary Value**: Total spending by the customer.
   - These metrics form the foundation of the **RFM (Recency, Frequency, Monetary)** model, which is often used in clustering for customer segmentation.
5. **Determining the Number of Clusters**
   - **Elbow Method**: Calculate the within-cluster sum of squares (WCSS) for different cluster values and identify the "elbow point" where the reduction rate of WCSS declines sharply.

- **Silhouette Score**: Measures how similar an object is to its own cluster compared to other clusters. A higher silhouette score indicates better-defined clusters.

6. **Applying Clustering Algorithm**

   - **Running the Algorithm**: Once the optimal number of clusters is determined, apply the chosen clustering algorithm to group customers based on their features.

   - **Analyzing Cluster Output**: Review the characteristics of each cluster. For instance, one cluster may represent high-frequency, high-spending customers, while another may indicate low-frequency, low-spending customers.

7. **Cluster Interpretation and Profiling**
   - Create profiles for each segment based on their defining characteristics. For instance:

     - **Cluster 1**: "Loyal High Spenders" – Customers with high purchase frequency and spending.

     - **Cluster 2**: "Infrequent Bargain Shoppers" – Customers with low frequency and low spending.

     - **Cluster 3**: "New and Potential Customers" – Recently joined customers with moderate engagement.

   - Use these profiles to design targeted marketing strategies, such as loyalty programs for high-value customers or discounts for infrequent shoppers.

Hypothetical Example

Suppose a retail company wants to segment its customers. After collecting and preprocessing the data, they apply K-Means clustering with an optimal kkk value of 4. The clusters generated are as follows:

- **Cluster 1**: High-frequency, high-value customers (Premium Customers)

- **Cluster 2**: Medium-frequency, medium-value customers (Regular Customers)
- **Cluster 3**: Low-frequency, high-value customers (Occasional Big Spenders)
- **Cluster 4**: Low-frequency, low-value customers (One-Time Shoppers)

Based on these segments, the company decides to:

- **Target Cluster 1** with exclusive rewards and personalized offers.
- **Engage Cluster 3** with seasonal promotions to increase frequency.
- **Re-engage Cluster 4** with discounts and product recommendations.

Benefits of Clustering in Customer Segmentation

- **Improved Marketing Efficiency**: Personalized campaigns based on customer segments lead to better conversion rates.
- **Enhanced Customer Experience**: By understanding customer preferences, the business can tailor experiences, resulting in higher satisfaction and loyalty.
- **Revenue Growth**: Targeted promotions for high-value segments can increase customer lifetime value.

Limitations and Considerations

- **Data Quality**: Reliable segmentation relies on accurate data; poor data can lead to misleading segments.
- **Dynamic Segments**: Customer behaviour may change over time, requiring periodic revaluation of segments.
- **Choosing the Right Algorithm**: Different clustering methods may yield different results.

Conclusion

Customer segmentation using clustering techniques allows businesses to create a more focused approach to customer management. By grouping customers based on similar characteristics, companies can drive engagement, optimize marketing efforts, and ultimately, foster customer loyalty and growth. In this case study, the company effectively used clustering to understand its customer base and strategize accordingly, illustrating the practical impact of data-driven customer segmentation.