Lookalike Model Report

Introduction

This report provides an overview of a lookalike model built to identify similar customers based on their purchasing behavior. By leveraging customer and transaction data, this model aims to enhance targeted marketing and customer relationship management.

Methodology

Data Merging

The first step involved merging customer, transaction, and product datasets:

- **Customer Data:** Contains demographic details of customers.
- **Transaction Data:** Records of customer transactions, including dates and total values.
- **Product Data:** Includes details about various products.

The merging process involved combining these datasets based on common keys (e.g., CustomerID and ProductID).

Data Transformation

To prepare the data for similarity analysis:

- **Pivoting Data:** The merged data was pivoted to create a user-product matrix, where rows represent customers, columns represent products, and values represent the total transaction value for each product by each customer.
- **Standardization:** The user-product matrix was standardized using StandardScaler to ensure consistent scaling across all features.

Model Building

A Nearest Neighbors model was employed to find similar customers:

- **Model Selection:** NearestNeighbors with cosine similarity was chosen to measure the distance between customers in the standardized user-product space.
- **Training:** The model was trained on the standardized user-product matrix.

Identifying Lookalikes

For each customer, the model identifies the closest neighbors (lookalikes):

- **Neighbor Retrieval:** For each customer, the model retrieves the top 4 nearest neighbors.
- **Similarity Calculation:** Distances between the target customer and their neighbors are converted to similarity scores (1 - distance).

Output Generation

The lookalike information is saved as a CSV file (Lookalike.csv) with the following structure:

- **CustomerID:** ID of the target customer.
- **Lookalikes:** List of similar customers and their respective similarity scores.

Conclusion

This lookalike model enables the identification of similar customers based on their purchasing behavior. By leveraging these insights, businesses can design targeted marketing campaigns, personalized recommendations, and effective customer retention strategies.