LynnsBaraka / recommendation-system

Code Issues Pull requests Actions Projects Wiki Security Insights

0 stars 0 forks 1 watching 9 Branches 0 Tags Activity

Public repository

master 9 Branches 0 Tags Go to file t Go to file Add file Code

LynnsBaraka ppt presentation now

.ipynb_checkpoints	Final changes on file	18 hours ago
.gitattributes	Adding files to git hub	last week
README.md	Update README.md	13 hours ago
data.csv	Added conclusions and cleaned file	20 hours ago
index.ipynb	Final changes on file	18 hours ago
recommendation system.pdf	ppt presentation	now
recommendation system.pptx	ppt presentation	now
shelf.jpeg	Added an image to the project description	5 days ago

README

Smart Retail Solutions Limited

Problem statement

A retail company aims to improve customer engagement and increase sales by implementing a Customer Segmentation & Recommendation System using Machine Learning (ML) techniques. The company wants to divide its customer base into distinct segments based on their purchasing behavior and preferences. Additionally, they want to develop a recommendation system that provides personalized product recommendations to customers within each segment.

Objectives

- Segment customers into meaningful groups based on their purchasing behavior, and other relevant factors.
- Develop a recommendation system that suggests products tailored to the preferences and interests of each customer segment.
- Enhance customer experience and increase sales by providing targeted and personalized recommendations.
- Evaluate the performance of the segmentation and recommendation system through metrics such as customer engagement, conversion rates, and revenue growth.

By addressing this business problem, the retail company can leverage machine learning technology to segment its customer base, deliver personalized recommendations, and drive business growth through enhanced customer engagement and sales.

Data Usage Strategy:

Through extensive analysis we were able to answer the following research questions:

- What are the key factors influencing customer purchasing behavior and preferences?

- How can customer segmentation techniques (clustering algorithms) effecϴvely divide the customer base into distinct segments?
- What are the characteristics and preferences of each customer segment, and how do they differ?
- How can content-based filtering recommendaϴons be integrated to provide more accurate and diverse recommendations to a customer based on their segment?
- How can the recommendation system address the cold start problem for new users or items with limited interacϴon history?

Procedure Overview:

Data Preparation:

Clean and preprocess the data, handling missing values, outliers, and inconsistencies.

Customer Segmentation:

Utilize K-means clustering algorithms to segment customers based on their purchasing behavior and preferences. Evaluate the clustering algorithms and determine the opϴmal number of segments using metrics such as silhoueΣe score or within-cluster sum of squares.

Recommendation System:

Develop a recommendation system that generates personalized product recommendations for each customer segment. Implement content-based filtering to suggest products tailored to individual preferences within each segment.

Model Evaluation:

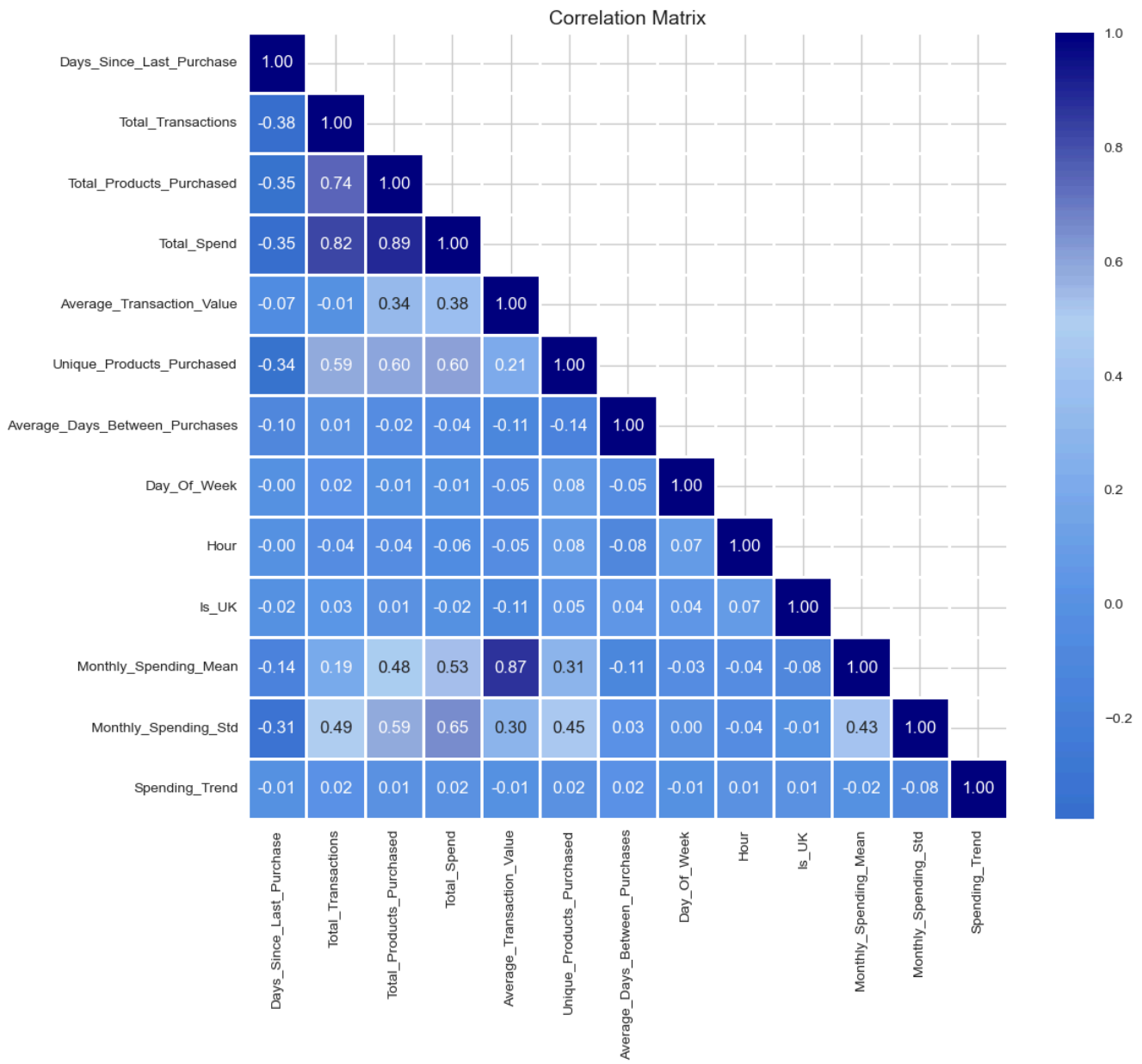
Use evaluation metrics and techniques to assess the effectiveness, relevance, and impact of your recommendation system in providing personalized and engaging recommendations to users.

Expectations of the Project

1. Improved Customer Engagement: The project aims to enhance customer engagement by providing personalized recommendations tailored to the preferences and interests of individual customers.
2. Increased Sales and Revenue: The ultimate goal of the project is to drive business growth by increasing sales and revenue. By delivering targeted and relevant product recommendations, the expectation is to boost conversion rates and encourage repeat purchases, resulting in higher transaction volumes and increased revenue for the company.
3. Enhanced Customer Experience: The project seeks to improve the overall customer experience by delivering a personalized and seamless shopping experience. The expectation is to provide customers with relevant product recommendations that match their preferences and needs, leading to higher levels of satisfaction and loyalty.
4. Effective Customer Segmentation: The project aims to segment the customer base into meaningful groups based on their purchasing behavior and preferences. The expectation is to identify distinct customer segments with unique characteristics and preferences, allowing the company to offer recommendations that meet the needs of each segment.

Correlation Analysis

In the below image we Looking at the heatmap, we can see that there are some pairs of variables that have high correlations:



This high correlations indicate that these variables move closely together, implying a degree of multicollinearity.can see that

Dimensionality Reduction(PCA)

In the below image the plot and the cumulative explained variance values indicate how much of the total variance in the dataset is captured by each principal component, as well as the cumulative variance explained by the first n components.

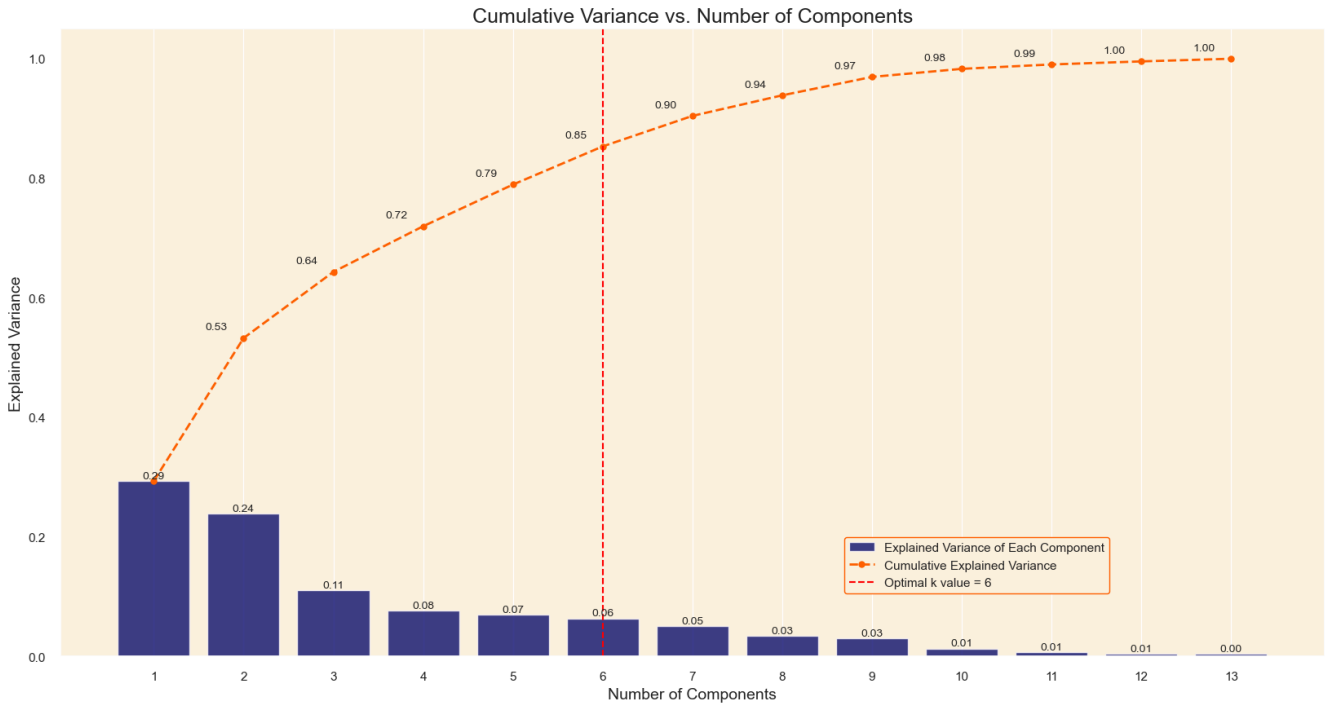
Here, we can observe that:

```
The first component explains approximately 28% of the variance.

The first two components together explain about 49% of the variance.

The first three components explain approximately 61% of the variance, and so on.
```

To choose the optimal number of components, we generally look for a point where adding another component doesn't significantly increase the cumulative explained variance, often referred to as the "elbow point" in the curve.

From the plot, we can see that the increase in cumulative variance starts to slow down after the 6th component (which captures about 81% of the total variance).

Considering the context of customer segmentation, we want to retain a sufficient amount of information to identify distinct customer groups effectively. Therefore, retaining the first 6 components might be a balanced choice, as they together explain a substantial portion of the total variance while reducing the dimensionality of the dataset.ge we get to see



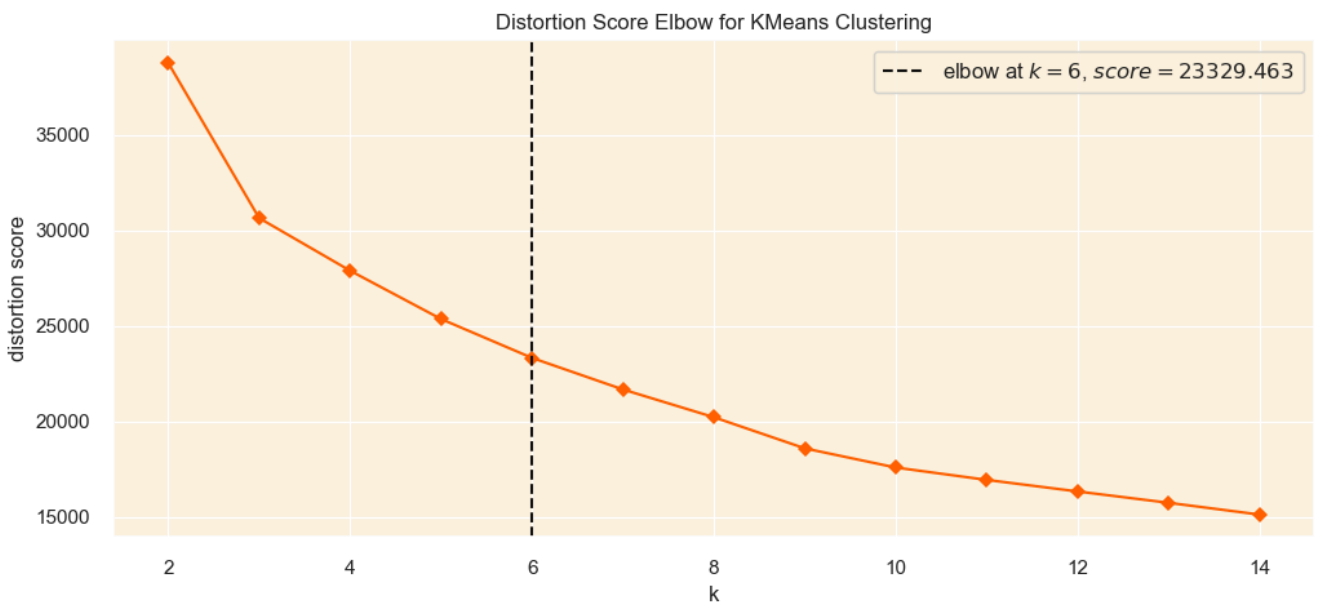
K-Means Clustering

KMeans algorithm, I will set the init parameter to k-means++ and n_init to 10. To determine the optimal number of clusters, I will employ the elbow method and silhouette analysis. Additionally, it might be beneficial to explore the use of alternative clustering algorithms such as GMM and DBSCAN in future analyses to potentially enhance the segmentation results.

Elbow Method

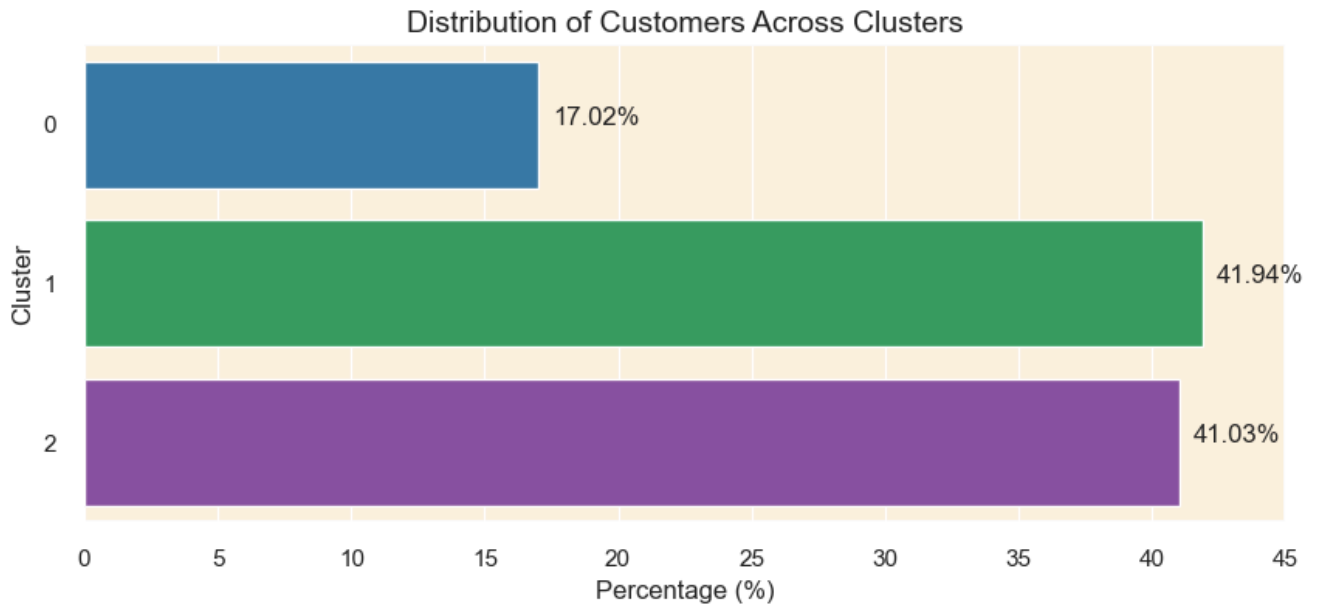
Optimal k Value: Elbow Method Insights

The optimal value of k for the KMeans clustering algorithm can be found at the elbow point. Using the YellowBrick library for the Elbow method, we observe that the suggested optimal k value is 6. However, we don't have a very distinct elbow point in this case, which is common in real-world data. From the plot, we can see that the inertia continues to decrease significantly up to k=6, indicating that the optimum value of k could be between 4 and 8. To choose the best k within this range, we can employ the silhouette analysis, another cluster quality evaluation method. Additionally, incorporating business insights can help determine a practical k value.

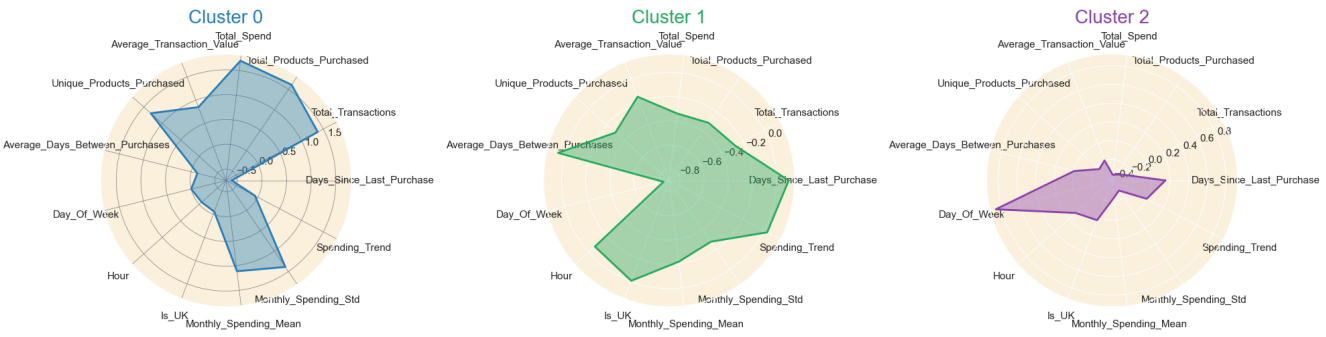


Clustering Evaluation

3D Visualization of Top Principal Components



Cluster Analysis and Profiling Using a Radar Chart Approach



- Cluster 0 has a high number of total transactions and total products purchased

- Cluster 1 has a high number of average days between purchases and days since last purchase.
- Cluster 2 has high number of the days of the week purchased.

RECOMMENDATION SYSTEM

Based on Frequently bought products:

Cluster 0 Recommendations:

1. JUMBO BAG RED RETROSPOT
2. WHITE HANGING HEART T-LIGHT HOLDER
3. LUNCH BAG RED RETROSPOT
4. REGENCY CAKESTAND 3 TIER
5. PARTY BUNTING
6. ASSORTED COLOUR BIRD ORNAMENT
7. SET OF 3 CAKE TINS PANTRY DESIGN
8. LUNCH BAG BLACK SKULL.
9. LUNCH BAG SUKI DESIGN
10. LUNCH BAG SPACEBOY DESIGN

Cluster 1 Recommendations:

1. WHITE HANGING HEART T-LIGHT HOLDER
2. REGENCY CAKESTAND 3 TIER
3. ASSORTED COLOUR BIRD ORNAMENT
4. PARTY BUNTING
5. REX CASH+CARRY JUMBO SHOPPER
6. POSTAGE
7. JUMBO BAG RED RETROSPOT
8. NATURAL SLATE HEART CHALKBOARD
9. PAPER CHAIN KIT 50'S CHRISTMAS
10. HEART OF WICKER SMALL

Cluster 2 Recommendations:

1. WHITE HANGING HEART T-LIGHT HOLDER
2. REGENCY CAKESTAND 3 TIER
3. ASSORTED COLOUR BIRD ORNAMENT
4. REX CASH+CARRY JUMBO SHOPPER
5. SET OF 3 CAKE TINS PANTRY DESIGN
6. HEART OF WICKER SMALL

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Contributors 4


LynnsBaraka

sammaash


BrantonKM


stephaniemwai

Languages

- **Jupyter Notebook** 100.0%