Diana Golman 7/20/2022

Capstone three project proposal: Customer segmentation

Can we identify customer clusters/ groups of a department store that have similar profiles and what characteristics do these customer clusters have?

Problem Identification:

Companies want to adjust their marketing campaigns, product promotions, and other individualized selling strategies. Identifying customer clusters/groups helps them to create targeted marketing and personalized services to the customers.

Context:
Identifying the need for a customer segmentation process is the first step to implementing a process that aligns with the overall business plan. When businesses don't have an effective customer segmentation process, they might find themselves providing the same service level for all customers and all products without focusing on the top-level customers or products that bring in the best margins.
The customer segmentation project should be continuously refined to help the business better respond to changing market trends. Companies' top-level customers should certainly receive more attention while building customer relationships and drawing retaining plans.
On the product level, without clear differentiation, businesses may invest in products that may not align with current market trends and profitability goals. Differentiation helps businesses to make more strategic investments in customer service and products. (https://blog.arkieva.com/customer-segmentation-strategy-issues/, By Editor, March 21st, 2017, Segmentation, Supply Chain Strategy)

Well-defined processes and technology can help support customer segmentation and the value it brings to a company. How the different segments are defined, and what the data in each customer segment is, is also essential to user /customer adoption and the successful usage of the segments in analytics. (FEBRUARY 6, 2019/BY MIKE WILSON, https://www.ironsidegroup.com/2019/02/06/challenges-customer-segmentation/)

This recognition and definition of each segment was previously a challenging and time-consuming task, that demanded hours of manually poring over different tables and querying the data in hopes of finding ways to group customers.
Machine learning algorithms can help marketing analysts find customer segments that would be very difficult to spot through intuition and manual examination of data.
(https://bdtechtalks.com/2020/12/28/machine-learning-customer-segmentation/)

With customers expecting brands to provide personalization, the brands who don't focus on providing that within a modern customer segmentation strategy will soon be missing out.

Diana Golman 7/20/2022

Behavioral data in combination with online users' behavior, demographics, and any other data available can be used in machine learning and significantly change the way a company perceives its customers, segmentation processes, and marketing or product strategies.
(https://formation.ai/blog/customer-segmentation-models-theres-a-better-approach-for-2022/
April 25, 2022)

Criteria for success :
The customer data will be used to segment and create clusters of customers. The company can use this segmentation to strategize its marketing or product strategies. More sophisticated, based on machine learning segmentation can lead to profit increase.

Scope of solution space:
A model will be created to segment the customer data and create clusters. The model will be evaluated for accuracy and prediction abilities. The best model will be applied in the industry as needed by a wide range of industries. The specific data set is closer to department stores and will be easily applied in similar industries.

Constraints: There are several potential challenges and constraints to this solution.:
The number of clusters based on the specific dataset can be not suitable to solve the problem. For example, a small number of clusters might not be able to provide a personalized enough service/product. On the other hand, if there are too many clusters there will be a problem with budget division or the result will not give us a clear overview of the clients/users.
Finally, the machine learning approach in general or this model specifically will be considered a change, and changes are difficult. Some marketing or analytics teams might be reluctant to incorporate untraditional practices that initially take more time. Of course in the long run it saves time and teams will have to be open to new practices.

Stakeholders: Marketing or Product analytics teams in any industry. The specific data set is closer to department stores and will be easily applied in similar industries.

Data source:
Dataset: Customer Personality Analysis (please click on the name) is a publicly available dataset on Kaggle. Acknowledgment: The dataset was provided by Dr. Omar Romero-Hernandez.
The dataset has 2240 entries that represent customers' interaction with the company's products.
It also has 29 features that can represent this information:

People

- ID: Customer's unique identifier
- Year_Birth: Customer's birth year
- Education: Customer's education level
- Marital_Status: Customer's marital status

- Income: Customer's yearly household income
- Kidhome: Number of children in customer's household
- Teenhome: Number of teenagers in customer's household
- Dt_Customer: Date of customer's enrollment with the company
- Recency: Number of days since customer's last purchase
- Complain: 1 if the customer complained in the last 2 years, 0 otherwise

Products

- MntWines: Amount spent on wine in the last 2 years
- MntFruits: Amount spent on fruits in the last 2 years
- MntMeatProducts: Amount spent on meat in the last 2 years
- MntFishProducts: Amount spent on fish in the last 2 years
- MntSweetProducts: Amount spent on sweets in last 2 years
- MntGoldProds: Amount spent on gold in the last 2 years

Promotion

- NumDealsPurchases: Number of purchases made with a discount
- AcceptedCmp1: 1 if customer accepted the offer in the 1st campaign, 0 otherwise
- AcceptedCmp2: 1 if customer accepted the offer in the 2nd campaign, 0 otherwise
- AcceptedCmp3: 1 if customer accepted the offer in the 3rd campaign, 0 otherwise
- AcceptedCmp4: 1 if customer accepted the offer in the 4th campaign, 0 otherwise
- AcceptedCmp5: 1 if customer accepted the offer in the 5th campaign, 0 otherwise
- Response: 1 if the customer accepted the offer in the last campaign, 0 otherwise

Place

- NumWebPurchases: Number of purchases made through the company’s website
- NumCatalogPurchases: Number of purchases made using a catalog
- NumStorePurchases: Number of purchases made directly in stores
- NumWebVisitsMonth: Number of visits to the company’s website in the last month

Deliverables:
A GitHub repo containing the project, a slide deck, and a project report.