IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Customer Segmentation for Banking Strategy using Machine Learning

Om Atre, Shrinil Modhave, Prasad Torane, S.B.Jadhav

Computer Department, MMCOE, Savitribai Phule Pune University

omatre2019.comp@mmcoe.edu.in

shrinilmodhave2018.comp@mmcoe.edu.in

prasadtorane2019.comp@mmcoe.edu.in

jadhavsb@mmcoe.edu.in

***Abstract*— Everything is becoming more mechanized in today's times, especially in this age of data. Data is the new oil. Many firms can benefit from this information in their marketing efforts. Every company wants to grow and be better than the competition, so they employ innovative ideas and invest in product development, but only to keep their customers and expand their market. Customers are segmented in today's commercial world. They analyze their consumers' behaviors and purchasing patterns and can assist them in generating profit. Customer segmentation refers to the process of categorizing a company's customers into groups based on their commonalities. Client segmentation is the process of determining how to interact with consumers in different groups to amplify the value of each customer to the company. There are many algorithms that can partition customers based on their behavior and other parameters.**

***Keywords*— -Card holders, Cluster analysis, machine learning, customer segmentation, private banking.**

I. INTRODUCTION

Customer segmentation is the process of classifying customers into specific groups based on shared characteristics. This allows companies to refine their messaging, sales strategies, and products to target, advertise, and sell to those audiences more effectively. This approach is used for both Business-to-Consumer (B2C) and Business-to-Business (B2B) marketing. B2C consumer segments are generally based on things like demographics, lifestyle, values, and needs. B2B marketers, on the other hand, tend to segment customers by industry, location, payment terms, or the specific products a company previously purchased.

Segmenting customers makes it easier for companies to sell their products and services. It is the difference between marketing a product to anyone and everyone, versus tailoring your messaging to address the needs or interests of specific groups of customers.

Once a company has established customer personas, it can map out a customer's journey that represents the decisions and steps a customer takes before, during, and after a transaction. This enables the bank to tailor its marketing messaging, branding, and pricing to every step in the journey.

Organizations often segment customers as part of a larger market research effort to understand the opportunities that exist in the marketplace.

You can use market research templates and other market research resources to understand why someone buys a product. But you also have to know who will buy the product. This is where customer segmentation comes into play. When you determine the characteristics of people who like and will buy your product, you can effectively divide your customer base into subgroups.

You will need to collect data to segment customers into specific groups. Surveys are a great way to get data related to demographics, purchasing patterns, preferences, and other distinct categories. You can also use interviews, existing customer data, focus groups, and other data collection methods[3]

II. MOTIVATION

Recent segmentation studies of bank customers and prospects provide insights into factors that influence bank selection. Acknowledging that there may be significant differences among your defined customer segments, these insights can help bank marketers fine-tune marketing planning in support of customer acquisition goals.

III. RELATED WORK

As of now, the types of segmentation methods are:

1. **Geographic Segmentation:**
   Geographic Segmentation is defined by where customers live and work or where businesses are placed. There are six factors that are used to perform segmentation namely:
   - Geographic Location (province, city, postal code, district, region)
   - Timezone
   - Climate
   - Cultural or Religious Preferences
   - Languages Spoken
   - Population Characteristics (urban, rural, population density)

   Geographic segmentation is useful in advertising, especially in US zip codes. Local businesses can easily advertise services by using specific zip codes.

2. **Demographic Segmentation:**
   Demographic Segmentation is mostly used to divide people into groups particularly based on their personal and lifestyle characteristics. Which includes:
   - Age
   - Gender Identity
   - Income
   - Ethnicity
   - Occupation
   - Religious Affiliation
   - Household Structure
   - Education

   Demographics are mostly used for targeting, especially for research surveys. People with similar likings or identities represent potential future customers, making it easier to evaluate marketing messaging.[2]

3. **Psychographic Segmentation:**
   Psychographic Segmentation gives importance to psychological features like personality, attitudes, and beliefs. Generating segments based on these features helps better grasp what consumers feel about your products.
   When executing market research, psychographic surveys are used to collect data to understand not just buying behaviors, but the incentive behind those behaviors.[2]

4. **Behavioral Segmentation:**
   Behavioral Segmentation is about grouping customers based on actions, rather than their psychology. It helps understand not just who they are but how they behave when they approach a certain transaction or do a purchase.

Zooming in on customers' habits, tendencies, and certain idiosyncrasies often involves analyzing certain purchasing behavior. What steps do they go through? How many one off purchases do they make? How many transactions do they make on average?

Four Factors which influence segmentation:

- Purchase Behavior
- Occasion and Timing
- Benefits Desired
- Customer Loyalty

It also further considers:

- Customers' stage in customer journey
- Their level of engagement with the product
- Their satisfaction level with the purchasing process

Through behavioral segmentation, one can understand what an end user experiences at the point of purchase, identify opportunities, evaluate risk if any, and design a strategy that optimizes and makes the customer experience more pleasant.[1,5]

5. **Technographic Segmentation:**

Technographic Segmentation identifies what systems and devices influence customers. For B2C segmentation, marketers must understand what technologies people use to purchase products and use those to enhance product marketing. People who consume products from a phone may differ from those who buy physically from a store. Each characteristic must be considered for marketing.

For B2B sales, segmentation can include what type of tech is used by customers, including software, automation tools for marketing purposes, frequently used softwares, and any other emerging technology. When combined with geography, industry, and personal information of consumers, technographic segmentation increases the effectiveness of reaching target audiences.

6. **Needs-Based Segmentation:**

Needs-Based Segmentation assists marketers refine their customers' segments by directing attention on how consumers get their psychological and practical needs met. Market segmentation surveys are beneficial for collecting this kind of qualitative information about consumers' needs.

Through surveys, consumers disclose which product features meet their requirements. Understanding these requirements and how they convey to customer pain points assists you create products that solve issues for particular customer segments.

Specific approaching, value propositions, and better engagement are all beneficial for needs-based segmentation. Panel surveys will disclose if companies are successfully fulfilling the needs not just once, but over a longer period of time.

7. **Value-Based Segmentation:**

Value-Based Segmentation emphasizes on what value customers get for the price they pay for a product or service. It has a direct influence on a product's price and what economic aid it yields. Clustering customers with similar economic features helps businesses understand the price sensitivity of a product or service. Companies can compare what price clients are willing to pay with their production costs to determine the future profitability of a product. Incorporating values-based segmentation with other segmentation methods will help us evaluate why people buy your products, what type of needs will be met, and what is the price customers are ready to pay for the value they will be receiving.

IV. SYSTEM ARCHITECTURE

The below figure shows the overall outline of our system. The first step consists of preprocessing of data. In this step, Data preprocessing is a process of preparing the unstructured data and making it suitable for a machine learning model. It is the first and very important step while creating a machine learning model.

A. Pre-processing of raw data

When creating a machine learning model, it is not always the case that we will have clean and formatted data. And before starting, it is mandatory to clean it and put it in a formatted way.[10] So, for this, we use data preprocessing.

The next step consisted of classifying customers in order to segment them into clusters. For this system we have used KMeans clustering[11], Principal Component Analysis (PCA).

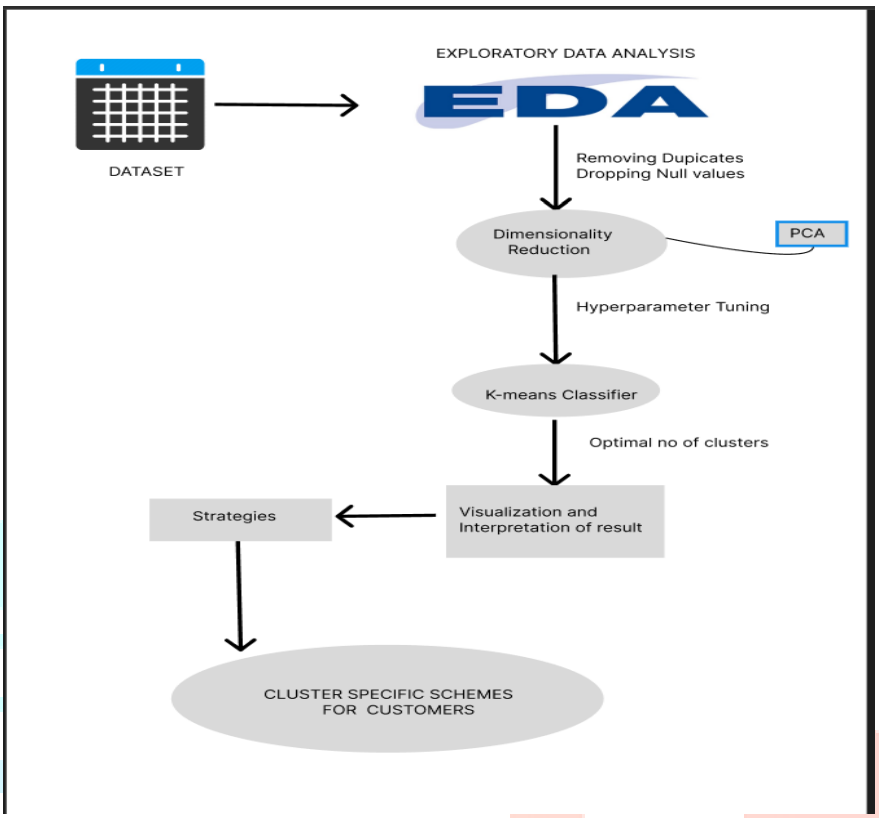

Fig. 1 shows the overall outline of our system.

B. Classify Customers:

When we try to understand the behavior of customers, initially we look at their history of interactions with the company, their attributes, metrics, scores, and the marketing feedback ratio.

There are a lot of sources and data that we can use to improve our knowledge and use it to make things better and improve our sales and invest our marketing budget for a higher ROI.

The next step consisted of classifying customers in order to segment them into clusters. For this system we have used KMeans clustering, Principal Component Analysis (PCA).

How the K-means algorithm works:

To process the data, the K-means algorithm in data starts with selecting a random group, a first group of centroids, which are used as the starting points for every cluster, and then performs repetitive calculations to optimize the positions of the cluster centroids.[4]

It stops generating and optimizing clusters when either:

- The centroids have secured — there is no change in their values because the clustering is successful.
- The defined number of iterations has been accomplished.

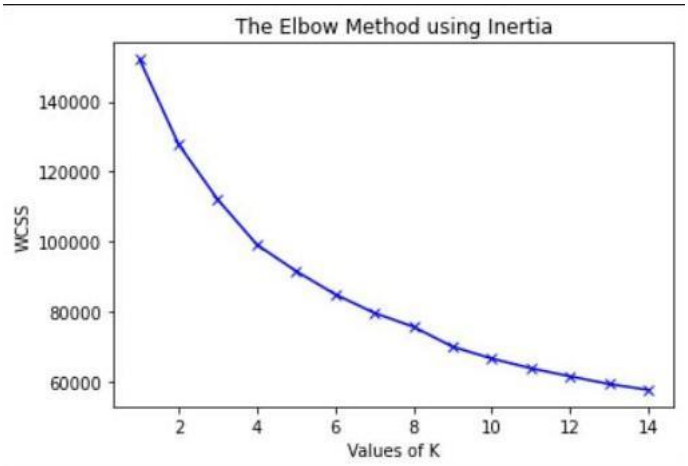

Fig. 2 The Elbow Method

Principal Component Analysis (PCA) - It is a statistical procedure that is a dimensionality reduction technique that converts a set of correlated variables to a set of uncorrelated variables. PCA however is the most widely used technique in exploratory data analysis and in machine learning for predictive or classification models. Moreover, PCA is an unsupervised technique used to determine the interrelations among a set of variables. Also known as a general factor analysis where regression determines a line that best fits.

Principal Component Analysis can also lead to low model performance after applying it if our dataset has a poor correlation or no correlation. The attributes need to be correlated to one other to apply PCA perfectly and efficiently. PCA provides us with a combination of features, and individual features significant from the original dataset are eradicated.[7] The principal axes with the most variance are the best suited principal components.

C.Proposed Work

After building the model we will be deploying it for further analysis. The Bank can use this clustered model to segment the new customers into different groups.[6] Based on their characteristics they can classified into various subgroups.[6,8] Later bank can analyse these characteristics and come up with different savings strategies or loan plans that would be beneficial for the customers. This efficient way of segmenting the customers and providing relevant strategies would attract more customers which will directly impact the growth of the bank.

VI.CONCLUSION

Our project will offer a quick, cost-effective way to create a solid customer segmentation strategy, target your customers, and get market insights for banks. Whether you're a big bank or a small one, you'll get a better understanding of who your customers really are, what they need, and how you can better serve them. We would have used K-Means Clustering[9] and created a Streamlit Application based on this clustering technique, where we take the customer details & identify which cluster the customer belongs to in order to give recommendations like saving plans, loans, portfolio management, etc. on target customer groups of that certain bank.

VII. REFERENCES

1. Customer Segmentation Using Machine Learning By Nikhil Patankar, Soham Dixit, Akshay Bhamare, Ashutosh Darpel, and Ritik Raina
2. E-Commerce Market Segmentation Based On The Antecedents Of Customer Satisfaction and Customer Retention By Brian Garda Muchardie, Anneta Gunawan, and Billy Aditya
3. Market Segmentation, Targeting and Positioning By Mark Anthony Camilleri
4. Application of K-Means Algorithm for Efficient Customer Segmentation: A Strategy for Targeted Customer Services By Chinedu Pascal Ezenkwu, Simeon Ozuomba, Constance Kalu
5. Review on Customer Segmentation Technique on Ecommerce By Juni Nurma Sari, Lukito Edi Nugroho, Ridi Ferdiana, and P. Insap Santosa
6. Segmentation of financial clients by attitudes and behavior: A comparison between Switzerland and Vietnam By Thuy Phan, Marc Oliver Rieger, and Mei Wang
7. The practical approach in Customers segmentation by using the K-Means Algorithm By E.Y.L Nandapala, and K.P.N Jayasena
8. Customer Segmentation using K-means Clustering By Tushar Kansal, Suraj Bahuguna, Vishal Singh, and Tanupriya Choudhury
9. Customer Segmentation Techniques on E-Commerce By Sumit Koul, and Trissa Merrin Philip
10. Credit users segmentation for improved customer relationship management in banking By Zita Bošnjak, and Olivera Grljevic
11. Single and Multiple point Spatial Queries Supporting Keywords for Searching Nearest Neighbors By Komal Chhajed, and Shailaja Jadhav