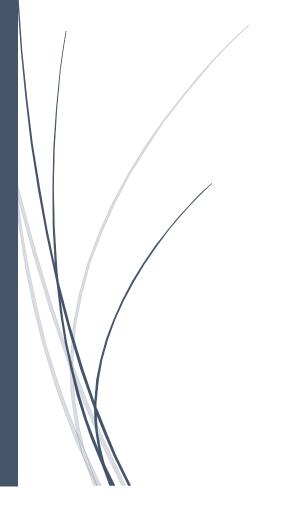
Mall Customer Segmentation

Team-2



Team Members:

Maneesh Madala - 19BCE7192 Patri Lalithya Manasa - 19BCD7103 Akella Siva Sai Atchyut - 19BCE7513

Jathin Kolla – 19BCD7192

Introduction:

Just like Sentiment Analysis is used to gain deeper insights into the customers' opinions and emotions about different products/services, Customer Segmentation is used for more targeted marketing. By categorizing the target audience into different buyer personas according to their needs, preferences, age, location, job, purchasing behavior, etc., brands can create customized products, marketing strategies, and offers/discounts, for a specific customer segment. This allows for higher customer satisfaction which eventually boosts the sales and revenue.

Customer Segmentation is one of the most extensively used applications of <u>unsupervised</u> <u>learning (ML)</u>.

Aim

Here we seek to achieve Value-based segmentation, where it differentiates customers by their economic value, grouping customers with the same value level into individual segments that can be distinctly targeted.

Market Needs

The marketers **must be able to relate their products to the** target segments. Market segmentation helps the marketers to understand the needs of the target audience and adopt specific marketing plans accordingly.

Data Information

The first step in the process of building the dataset is to collect pertinent data that can represent the customer spending score and Annual Income

This dataset is composed by the following five features:

Customer ID: Unique ID assigned to the customer

Gender: Gender of the customer

Age: Age of the customer

Annual Income (k\$): Annual Income of the customer

Spending Score (1-100): Score assigned by the mall based on customer behavior and spending nature.

In this dataset we have 1500 samples to study.

Data Analysis

The technique of customer segmentation is dependent on several key differentiators that divide customers into groups to be targeted. Data related to demographics, geography, economic status as well as behavioural patterns play a crucial role in determining the company direction towards addressing the various segments

Data Pre-processing:

After data is collected, it is important to lay the ground for data analysis by applying various pre-processing operations to address potential imperfections in the raw collected data. Data Collection methods and conditions are not flawless, especially in the context of Remote Sensing data where faulty equipment or human's inattention can lead to a noisy dataset containing errors, redundancies, and outliers.

Data Mining Algorithms:

In this project, you will use the K-means algorithm for clustering an unlabelled dataset. The K-means clustering algorithm can effectively visualize the age and gender distributions in the dataset. Further, it will also analyse annual incomes and spending patterns.

Conclusion

- K-Means Clustering is a powerful technique in order to achieve a decent customer segmentation.
- Customer segmentation is a good way to understand the behaviour of different customers and plan a good marketing strategy accordingly.
- There isn't much difference between the spending score of women and men, which leads us to think that our behaviour when it comes to shopping is pretty similar.
- Observing the clustering graphic, it can be clearly observed that the ones who spend more money in malls are young people. That is to say they are the main target when it comes to marketing, so doing deeper studies about what they are interested in may lead to higher profits.
- Although younglings seem to be the ones spending the most, we can't forget there are more people we have to consider, like people who belong to "middle class" and it seems to be the biggest cluster.

• Promoting discounts on some shops can be something of interest to those who don't actually spend a lot and they may end up spending more!

References

- [1] I. S. Dhillon and D. M. Modha, "Concept decompositions for large sparse text data using clustering," Machine Learning, vol. 42, issue 1, pp. 143-175, 2001.
- [2] S. Dasgupta and Y. Freund, "Random Trees for Vector Quantization," IEEE Trans. on Information Theory, vol. 55, pp. 3229-3242, 2009.
- [3] https://data-flair.training/blogs/r-data-science-project-customer-segmentation/
- [4] https://www.kaggle.com/vjchoudhary7/customer-segmentation-tutorial-in-python