CUSTOMER SEGMENTATION USING K-MEANS CLUSTERING

SHEERLA NEHA(21RG1A05C0)
POPULLA VASUNDHARA(21RG1A05B3)
PATLOLLA PRATHUSHA(21RG1A05B2)
 RODDA RUCHITHA(21RG1A05C0)

Under the guidance of
Mrs.K.SWETHA

ABSTRACT:

–>The spirit of the modern age is characterized by innovation, with everyone engaged in a race to outshine the competition.

–>Today's businesses thrive on such innovation, captivating customers with their products. However, the vast array of products available often leaves customers bewildered about what to purchase, while companies struggle to identify which customer segments to target.

–>This is where machine learning becomes crucial. Various algorithms are applied to uncover hidden patterns in data, facilitating better decision–making for the future. –>This elusive concept of which segment to target becomes clear through segmentation.

–> Customer segmentation involves grouping customers with similar behaviors into the same segment and those with different patterns into separate segments

–> A Python program was developed and trained using a standard scaler on a dataset with two features from 200 training samples from a local retail shop.

–> The features include the average amount spent by customers and the average number of visits to the shop annually.

–>By applying clustering, five customer segments were identified: Careless, Careful, ––Standard, Target, and Sensible customers.

INTRODUCTION:

-> As new businesses continue to emerge daily, it has become crucial for established companies to adopt marketing strategies to remain competitive in a cutthroat market.

->With the ever-growing customer base, it has become increasingly challenging for companies to meet the diverse needs of all their customers.

-> This is where data mining plays a pivotal role in uncovering hidden patterns within the company's data repositories. One significant application of data mining is customer segmentation, which groups customers with similar behaviors into distinct clusters, thereby simplifying the management of a large customer base.

->Customer segmentation can greatly influence marketing strategies by opening up numerous new avenues, such as identifying the most suitable product for each segment, ->customizing marketing plans for different segments, offering targeted discounts, and uncovering previously unknown relationships between customers and products.

->Clustering has proven to be an effective method for implementing customer segmentation. As an unsupervised learning technique, clustering can identify groups within unlabelled datasets.

->In this study, three different clustering algorithms were applied to a dataset with two features and 200 records to demonstrate their effectiveness in customer segmentation.

GRAPHICAL REPRESENTATION OF COMPARISON OF DIFFERENT ALGORITHMS:

- Silhouette Score:
->It is a way of measuring how well the data point has been clustered into the correct cluster.
->Figure above displays the silhouette score for the three algorithms applied in this paper, the graph shows there is not much significant difference in K-means and Agglomerative clustering. Hence, these two algorithms were able to cluster our data well than Mean shift algorithm as displayed by the low value of silhouette score.
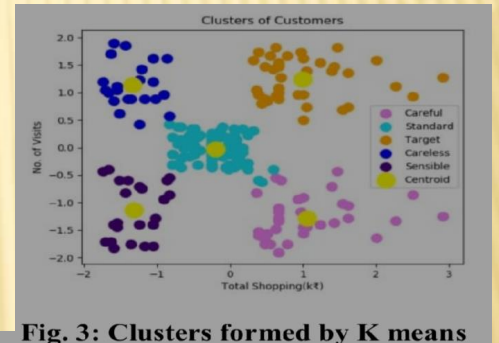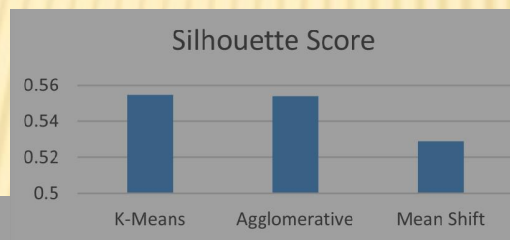

Fig. 3: Clusters formed by K means

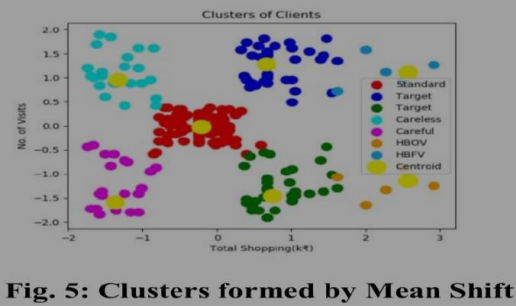Fig. 6: Comparison of Silhouette Score
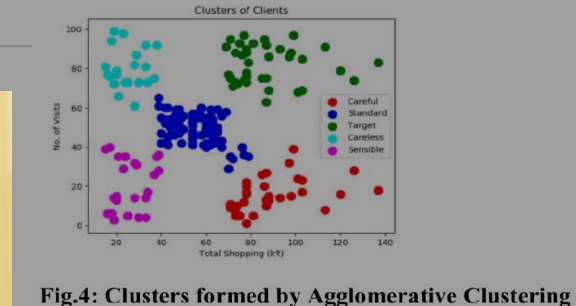

Fig. 5: Clusters formed by Mean Shift

Fig.4: Clusters formed by Agglomerative Clustering

ANALYSIS OF EXISTING SYSTEM:

Customer segmentation is a critical practice in marketing and business strategy, aiming to divide a customer base into groups with similar characteristics, behaviors, or needs. Two common methods for customer segmentation are earlier methods presumably traditional or heuristic-based.

DISADVANTAGES:

->May oversimplify customer behaviors or miss important nuances. Segments can be broad and less precise.

->Doesn't capture diverse behavioral patterns beyond transactional history.

->Requires robust data and sophisticated analytics to derive meaningful segments.

->Many earlier segmentation systems are static and do not adapt well to changes in customer behaviors or market dynamics over time. This rigidity can lead to ineffective marketing strategies and missed opportunities for personalized customer engagement.

->As businesses grow and collect more data, traditional segmentation methods may struggle to scale effectively. They may not be able to handle large volumes of data or efficiently process the complexity of customer interactions across multiple channels.

ANALYSIS OF PROPOSED SYSTEM:

customer segmentation using K-means clustering offers businesses a robust method to uncover actionable insights, enhance customer understanding, and drive targeted marketing efforts

When comparing current segmentation using K-means clustering to earlier methods such as heuristic or traditional segmentation approaches, several significant differences and advantages can be identified

ADVANTAGES:

->K-means Clustering: Utilizes mathematical algorithms to identify clusters based on data patterns, allowing for more precise segmentation. It can uncover subtle variations and group customers into distinct clusters based on multiple dimensions

->Analyses multiple variables simultaneously, incorporating complex interactions between variables to form clusters. It can handle large datasets and various types of data (numeric, categorical) effectively.

->Offers flexibility to adapt and update clusters in real-time as new data becomes available. This dynamic approach allows businesses to respond quickly to changing customer behaviors and market conditions.

->Scales well with larger datasets and can handle complex segmentation tasks efficiently using computational algorithms.

CONCLUSION:

->Customer segmentation through K-means clustering is a robust method that categorizes customers based on similarities in behavior, preferences, or characteristics.

->By grouping customers into distinct clusters, businesses can effectively target their marketing efforts, tailor products or services to meet specific segment needs, and allocate resources more efficiently.

->This approach enhances customer satisfaction by delivering personalized experiences that resonate with each segment's unique traits.

->However, it's crucial to recognize the limitations of K-means clustering, such as its assumption of spherical clusters and equal variance, which may not always align perfectly with real-world customer data.

->Businesses should validate clusters and supplement quantitative findings with qualitative insights to ensure accurate segmentation and meaningful strategic decisions.

->Continuous monitoring and adaptation of segmentation strategies are essential to reflect evolving customer behaviors and market dynamics, ensuring that segmentation remains relevant and beneficial over time.

->Overall, K-means clustering provides a valuable framework for customer segmentation, fostering -deeper customer insights and driving sustainable business growth.