1. **Abstract**

Customer segmentation is a strategic approach that has the potential to significantly boost revenue for your mall. To enhance the mall's profitability, the authorities recognize the need to attract and engage customers more effectively. However, traditional advertising methods often yield limited results, with only a 30-40% conversion rate. To address this challenge, the mall is embracing the power of customer segmentation, a widely adopted application of unsupervised learning. This shift in strategy is expected to drive a substantial increase in conversion rates, aiming for an impressive 90-95%. Our project centers around the application of the K-means clustering algorithm to group customers based on two critical factors: their Spending Score and Annual Income. By identifying and understanding these distinct customer clusters, we can tailor our marketing campaigns, promotions, and product offerings to resonate with each group's unique preferences and needs. The main aim of our mall customer segmentation project using the K-means algorithm represents a forward-thinking strategy to optimize marketing efforts and attract potential customers with the highest likelihood of making purchases.

1. **Introduction**
2. **Introduction**

Over the years, the competition amongst businesses is increased and the large historical data that is available has resulted in the widespread use of data mining techniques in extracting the meaningful and strategic information from the database of the organization. Data mining is the process where methods are applied to extract data patterns in order to present it in the human readable format which can be used for the purpose of decision support. According to, Clustering techniques consider data tuples as objects.

They partition the data objects into groups or clusters, so that objects within a cluster are similar to one another and dissimilar to objects in other clusters. Customer Segmentation is the process of division of customer base into several groups called as customer segments. The segmentation is based on the similarity in different ways that are relevant to marketing such as gender, age, interests, and miscellaneous spending habits. The customer segmentation has the importance as it includes, the ability to modify the programs of market so that it is suitable to each of the customer segment, support in business decision; identification of products associated with each customer segment and to manage the demand and supply of that product.

1. **Problem Statement**

Customer Segmentation is the best application of unsupervised learning. Using clustering, identify segments of customers in the dataset to target the potential user base. They divide customers into various groups according to common characteristics like gender, age, interest, and spending habits so they can market to each group effectively. Use K-Means Clustering and also visualize the gender and age distributions. Then analyze their annual income and spending scores. As it describes about how we can divide the customers based on their similar characteristics according to their needs by using k-means clustering which is a classification of unsupervised machine learning.

**2.1 SCOPE**

The scope of our Project Idea is:

1] Gather comprehensive data on customer spending behavior and annual income. This data will serve as the foundation for our customer segmentation analysis.

2] Select and engineer relevant features that capture customer behavior, income levels, and spending patterns to ensure the accuracy and effectiveness of our clustering algorithm.

3] Utilizing the K-means algorithm, we will cluster customers into distinct groups based on their spending scores and annual income.

4] This will help us identify clusters of high-income customers with low spending scores.

5] Determine potential customers within low-spending segments with high income.

6] Optimize advertising campaigns to target potential customers effectively.

**2.2 SOFTWARE AND HARDWARE ANALYSIS**

**2.3 REQUIRMENT ANALYSIS**

**2.4 LIBRARIES / PACKAGES USED**

1. **DATASET AND SOURCE CODE**

**A. Description of Dataset**

The dataset 'Mall\_Customers.csv' serves as the foundation for our customer segmentation project aimed at increasing mall revenue. It encompasses crucial information about mall visitors, facilitating a data-driven approach to enhance customer engagement and profitability.The dataset comprises five columns, each providing valuable insights into the mall's customer base:

1] CustomerID: A unique identifier assigned to each mall customer. This identifier enables individual tracking and analysis.

2] Gender: This categorical attribute denotes the gender of the customers, distinguishing between male and female shoppers. It allows for gender-based segmentation analysis.

3] Age: A numeric feature representing the age of each customer. Age is a key factor influencing shopping behavior and preferences.

4] Annual Income (k$): Another numeric attribute, this one signifies the annual income of customers in thousands of dollars. Income levels have a significant impact on purchasing power.

5] Spending Score (1-100): This numeric attribute quantifies the spending behavior of customers on a scale from 1 to 100.

**B. Datatypes:**

CustomerID: Numeric (Unique Identifier)

Gender: Categorical (Male or Female)

Age: Numeric (Continuous)

Annual Income (k$): Numeric (Continuous)

Spending Score (1-100): Numeric (Continuous)

**C. Link Of Dataset & Code:**

<https://github.com/PrasadThorve/MLMiniProject>

1. **OUTPUT**

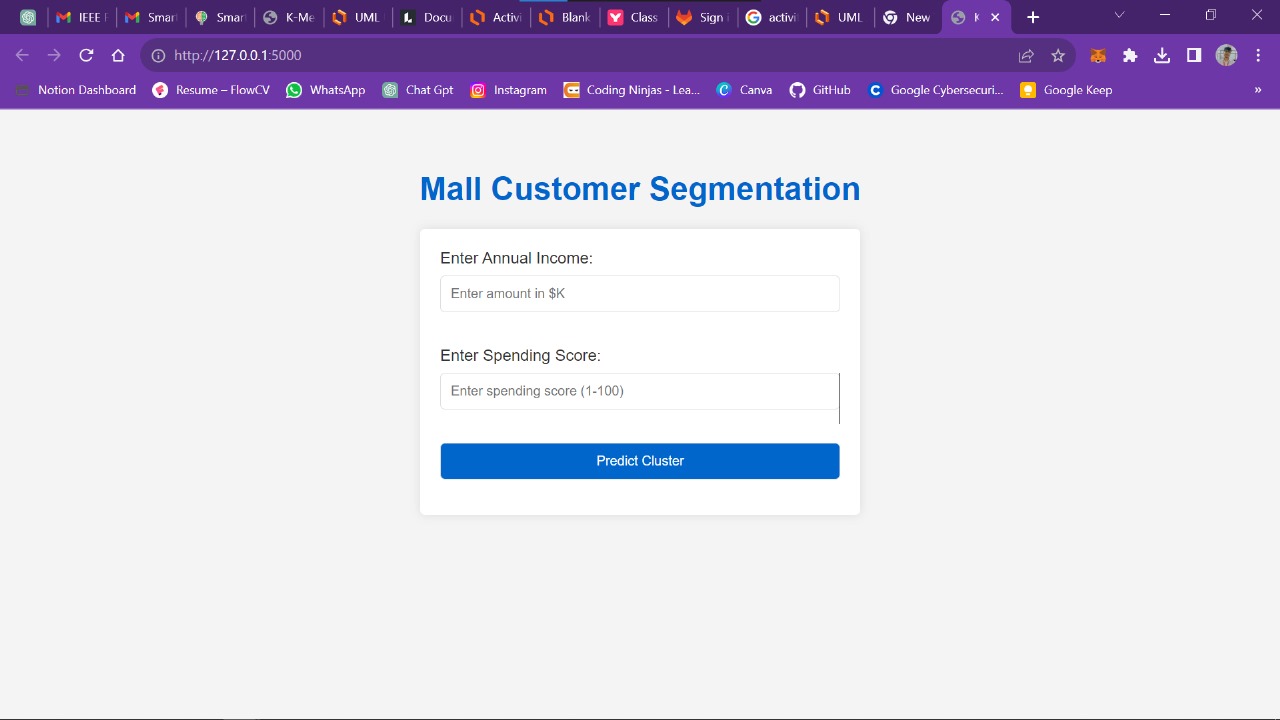


Fig.:-1 User Interface of Mall Customer Segmentation

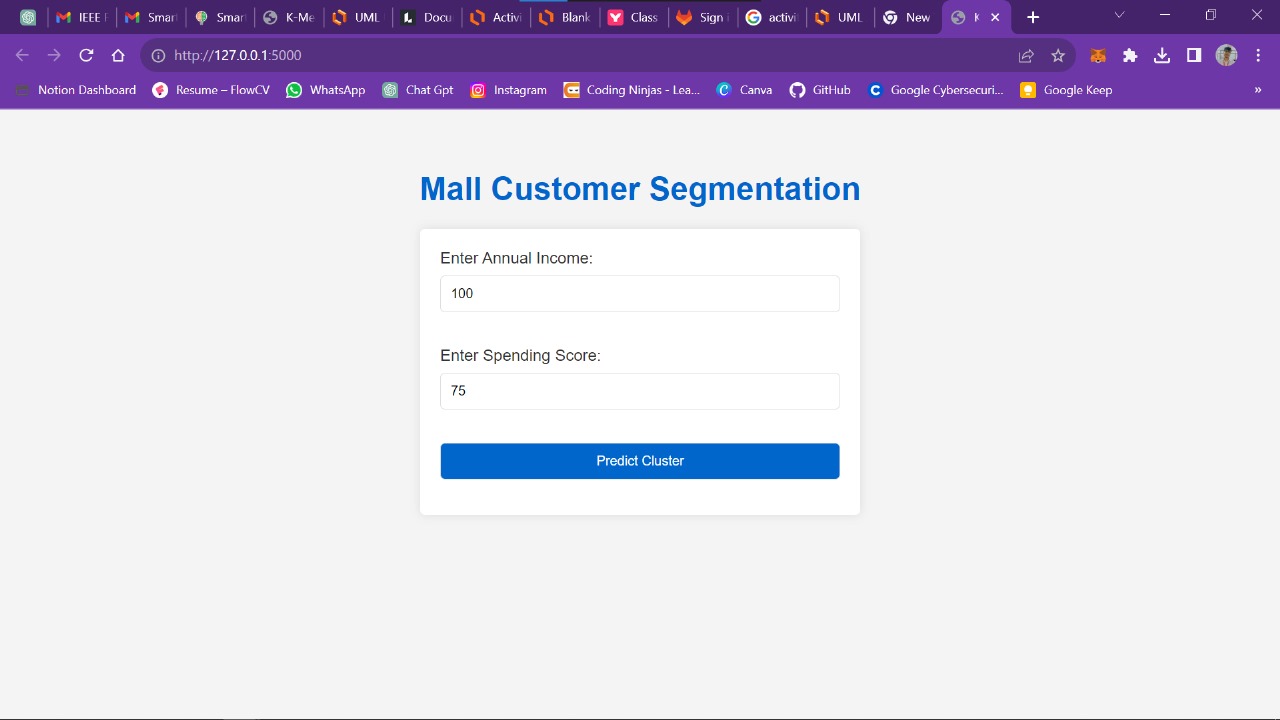


Fig.:-2 Input for predicting the cluster

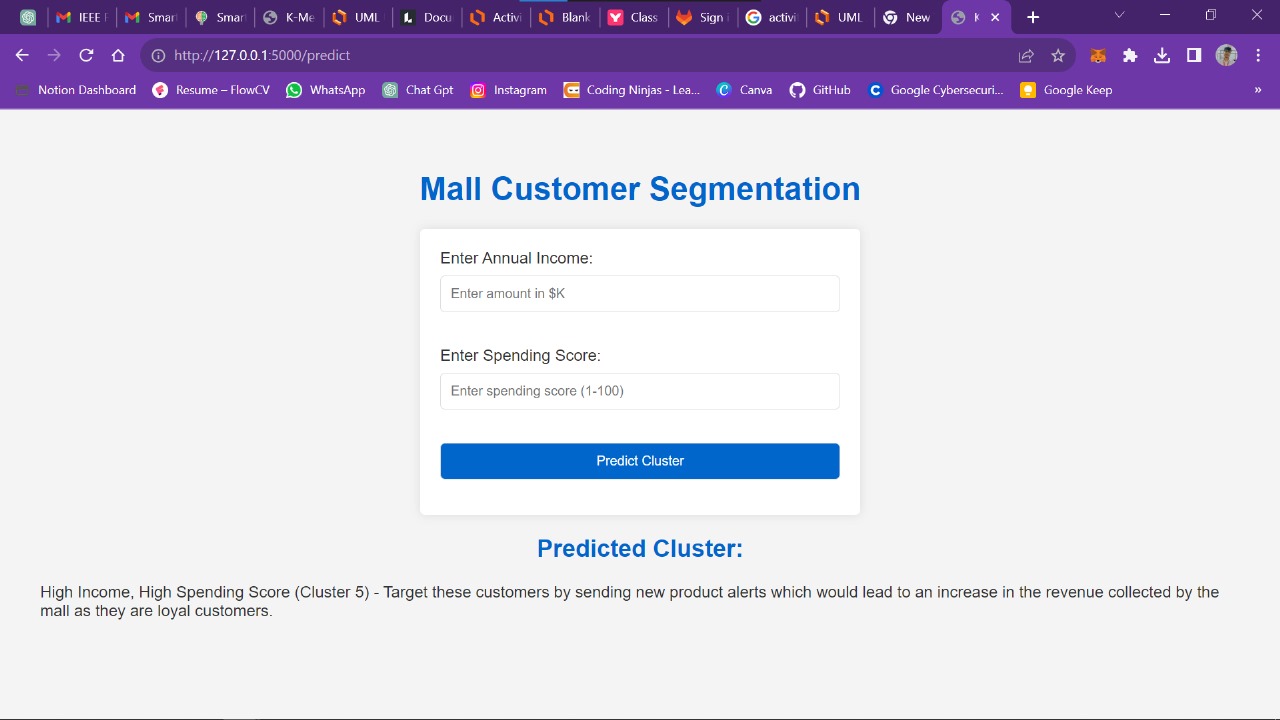


Fig.:- 3 Final Output of predicted cluster

1. **CONCLUSION**

So, we concluded that the , 1) The Highest income , high spending can be targeting these type of customers as they earn more money and spend as much as they want. 2) Highest income, low spending can be targeting these types of customers by asking feedback and advertising the product in a better way. 3) Average income, Average spending may or may not be beneficial to the mall owners of this type of customers. 4) Low income, High spending can be targeting these types of customers by providing them with low-cost EMI’s etc. 5) Low income, Low spending doesn’t target these types of customers because they earn a bit and spend some amount of money. So high income, high spending are the most beneficial ones to the mall owners which increases the owner’s business. Using market segmentation, companies are able to identify their target audiences and personalize marketing campaigns more effectively. This is why market segmentation is key to staying competitive. It allows you to understand your customers, anticipate their needs, and seize growth opportunities.