Mall Customer Segmentation Data

Abhinav Duvvuru   
 Artificial Intelligence and Machine Learning  
Lambton CollegeToronto, Canada  
abhinavduvvuru@gmail.com  
  
Sandeep Reddy Alla  
Artificial Intelligence and Machine LearningLambton CollegeToronto, Canada  
itsmealla05@gmail.com  
  
SivaPriya TatiReddy  
Artificial Intelligence and Machine LearningLambton CollegeToronto, Canada  
sivapriya721@gmail.com   
  
Uma Maheswararao Gutta Artificial Intelligence and Machine LearningLambton CollegeToronto, Canada umamaheshgutta@gmail.com

*Abstract*— In a study, we analyze the mall customer data and divide the customers into different groups. Now a days customers are the main source of income for Malls or shopping complexes. Shopping complexes use their customers data and develop ML models to target their audience which finally increases sales.

# Introduction

To increase the business and profit enhancement deal with customer data is more important to segment the data and group the customers based on the money spent ,behavior Which helps the marketing team to improve their sales.

# Subject of the study

Mall Customer data is an interesting dataset that has hypothetical customer data. The dataset used in this project is the data obtained from Kaggle, which has Customer ID, Customer Gender ,Customer Age, Annual Income of the customer, Spending score of the customer. The dataset has total 5 attributes. Below are the list of attributes and their type:

1. Customer ID (numerical) - ID of the Customer.

2. Customer Gender(categorical) -Gender of the customer.

3. Customer Age – Age of the customer.

4. Annual Income – Annual income of the customer

5. Spending score - Score assigned by the mall based on customer behavior and spending nature

# data extraction and cleaniing

The data set which we use cannot be implemented to the model without proper cleaning and extraction. We plan to apply sK-Means clustering algorithm to cluster the data .We do data preprocessing steps like checking for Null values or Missing values.

# model building

After converting and successfully processing data. We use elbow method or PCA(principal component analysis) to find the number of clusters we have chosen K-means clustering unsupervised learning algorithm this is the best algorithm best fits the customer dataset to understand customer data.

# Visualizing The results

Here We try to visualize the data based on wide variety of ages and correlations between the income and spending scores.visualize the patterns and customers into groups based on the k-value.

# Conclusion

Customer segmentation is a separation of a market into multiple distinct groups of consumers who share the similar characteristics. Segmentation of market is an effective way to define and meet customer needs. Unsupervised Machine Learning Techniques, K-Means Clustering Algorithm, Minibatch K-Means and Hierarchical Clustering are used to perform Analysis is carried out to predict the target customers who can be easily converged, among all the customers. In order to allow the marketing team to plan the strategy to market the new products to the target customers which are similar to their interests.

##### References

1. “Customer segmentation Dataset”. Kaggle [online dataset], November 2021.Available:Kaggle,https://www.kaggle.com/vjchoudhary7/customer-segmentation-tutorial-in-python [Accessed: November 16, 2021]