Intern Name: Muskaan
Intern ID: 170007

PROJECT ACTIVITY - 4

Mall Customer Segmentation using K-Means clustering

Market segmentation is the activity of dividing a broad consumer or business market, normally consisting of existing and potential customers, into sub-groups of consumers based on some type of shared characteristics.

Companies employing customer segmentation operate under the fact that every customer is different and that their marketing efforts would be better served if they target specific, smaller groups with messages that those consumers would find relevant and lead them to buy something. Companies also hope to gain a deeper understanding of their customers' preferences and needs with the idea of discovering what each segment finds most valuable to more accurately tailor marketing materials toward that segment.

Malls or shopping complexes are often indulged in the race to increase their customers and hence making huge profits. To achieve this task machine learning is being applied by many stores already. It is amazing to realize the fact that how machine learning can aid in such ambitions. The shopping complexes make use of their customers' data and develop ML models to target the right ones. This not only increases sales but also makes the complexes efficient.

Problem Statement

You own the mall and want to understand the customers like who can be easily converge [Target Customers] so that the sense can be given to marketing team and plan the strategy accordingly.

Download Dataset from Here: Mall_Customers.csv

Dataset overview:

1. CustomerID: It is the unique ID given to a customer
2. Gender: Gender of the customer
3. Age: The age of the customer
4. Annual Income(k$): It is the annual income of the customer
5. Spending Score: It is the score (out of 100) given to a customer by the mall authorities, based on the money spent and the behavior of the customer.

Let's analyze data from mall customers to dig intuitions about costumer convergence and apply K-Means algorithm to develop insights about how to deal with different clusters.

Steps to solve the problem:

1. Importing Libraries.

2. Exploration of data.

3. Data Visualization.

   - Count Plot of Gender
   - Plotting the Relation between Age, Annual Income and Spending Score
   - Distribution of values in Age, Annual Income and Spending Score according to Gender

Intern Name: Muskaan
Intern ID: 170007

4. Clustering using K-Means.

    - Segmentation using Age and Spending Score
    - Segmentation using Annual Income and Spending Score
    - Segmentation using Age, Annual Income and Spending Score

5. Selection of Clusters.

6. Plotting the Cluster Boundary and Clusters.

7. 3D Plot of Clusters.

Conclusions:

1. How to use this knowledge?

Instructions:
1. Submission on GitHub and link for your GitHub on Moodle.
2. Coding conventions are properly used.

Solution:

https://github.com/MuskaanSehgal/BDDA2_ProjectActivity4.git