TEAM- 04

UNSUPERVISED LEARNING

**MODEL USED: CATEGORIZING CREDIT CARD CUSTOMERS**

**SUBMITTED BY:**

Talasila Balaji -balajitalasila2@gmail.com

Arunkumar M -arunkumar777ms@gmail.com

Melvin R -melvinvmj@gmail.com

Reenarao R M -reenamurugan1728@gmail.com

Muli Likhith Reddy -likhithreddymuli@gmail.com

Mopuri Rishitha -rishimopuri@gmail.com

**ABSTRACT:**

This project involves clustering credit card customers based on their usage patterns and credit limits using the K-means algorithm. The dataset used in this project represents a sample of 660 credit card customers, with key attributes such as average credit limit, total number of credit cards, frequency of bank visits, online visits, and call made to customer service. The elbow method was employed to determine the optimal number of clusters based on its underlying structure of data. Once the optimal number of clusters was identified, the K-means Algorithm was applied, segmenting the customers into three distinct groups. These clusters provide insights into the different types of customers based on their credit limits and interaction with banking services. For instance, one cluster might represent high credit limit customers who rarely visit the bank, while another could represent customers with multiple credit cards and frequent customer service interactions. The results were visualized through a scatter plot, focusing on the relationship between average credit limit and the number of credit cards, which revealed clear distinctions between the clusters. The clustered data was then saved to a new CSV file for further analysis. This clustering model can be leveraged for targeted marketing strategies, personalized customer service, and better financial management by identifying distinct customer segments and understanding their behavior. This allows businesses to optimize their offerings and improve customer satisfaction by catering to the specific needs of each segment.

**DESCRIPTION OF DATASET:**

This dataset contains information on credit card customers, focusing on their credit usage and interactions with banking services. The data provides insights into customer behavior, which can be useful for segmentation, targeted marketing, and customer relationship management. Here’s a detailed description of the key features in the dataset:

Key Features:

1. **Sl\_No**: A unique serial number assigned to each record in the dataset.

2. **Customer Key**: A unique identifier for each customer.

3. **Avg\_Credit\_Limit**: The average credit limit assigned to the customer across all their credit cards.

4. **Total\_Credit\_Cards**: The total number of credit cards owned by the customer.

5. **Total\_visits\_bank**: The total number of times the customer has visited a bank branch.

6. **Total\_visits\_online**: The total number of times the customer has accessed the bank’s online portal or website.

7. **Total\_calls\_made**: The total number of calls made by the customer to the bank’s customer service.

This dataset offers a comprehensive view of credit card customers, enabling businesses to make data-driven decisions to enhance customer satisfaction and optimize service offerings.

**ALGORITHM USED:**

In this project, the K-Means clustering algorithm is used to segment credit card customers based on their credit usage and interaction with banking services. K-Means is a popular unsupervised machine learning algorithm that groups data points into clusters based on their similarities.

**ALGORITHM DETAILS:**

1. import libraries pd, KMeans, StandardScaler, plt
2. Load the data from csv file
3. Select the relevant columns for clustering
4. Standardize the features
5. Determine the optimal number of clusters using the elbow method
6. Plot the elbow graph
7. According to the elbow plot the optimum number of clusters is 3
8. Save the clustered data to a new CSV file
9. Plot the clustered data

**OUTPUT:**

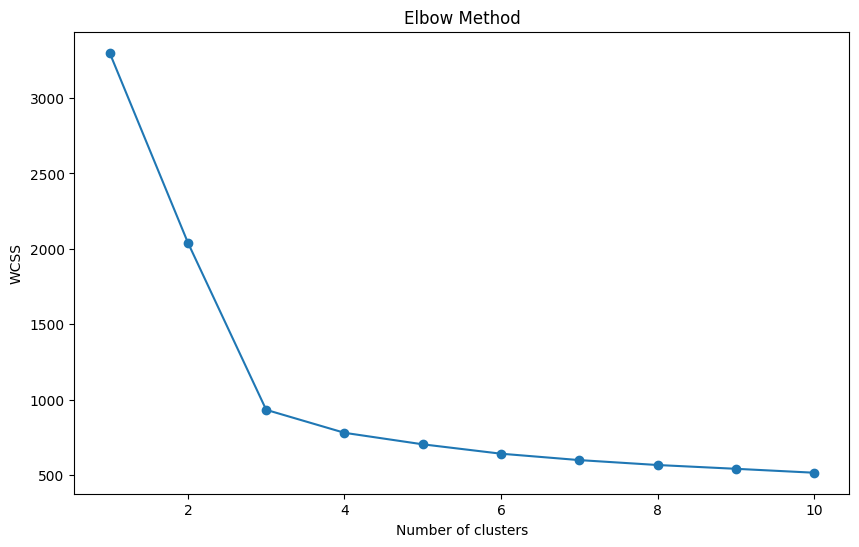


Fig 1: output after performing elbow method

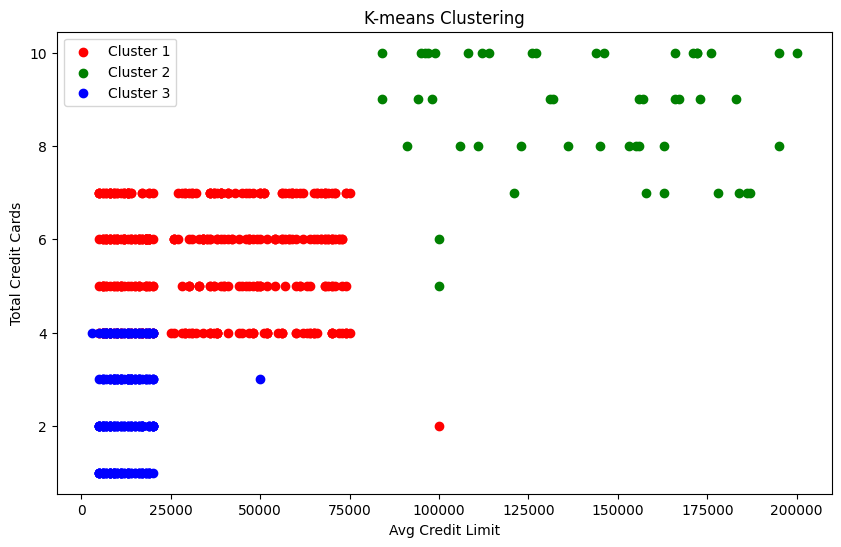


Fig 2: output after clustering completed and saved to 'clustered\_data.csv'

**CONCLUSION:**

The K-means clustering divided the dataset into three distinct groups based on customer behavior. The clustered data has been saved to a new CSV file named clustered\_data.csv. The visualization provides insights into how customers differ in terms of their credit limit and number of credit cards, which can inform strategic decisions.