Mall Customer Segmentation

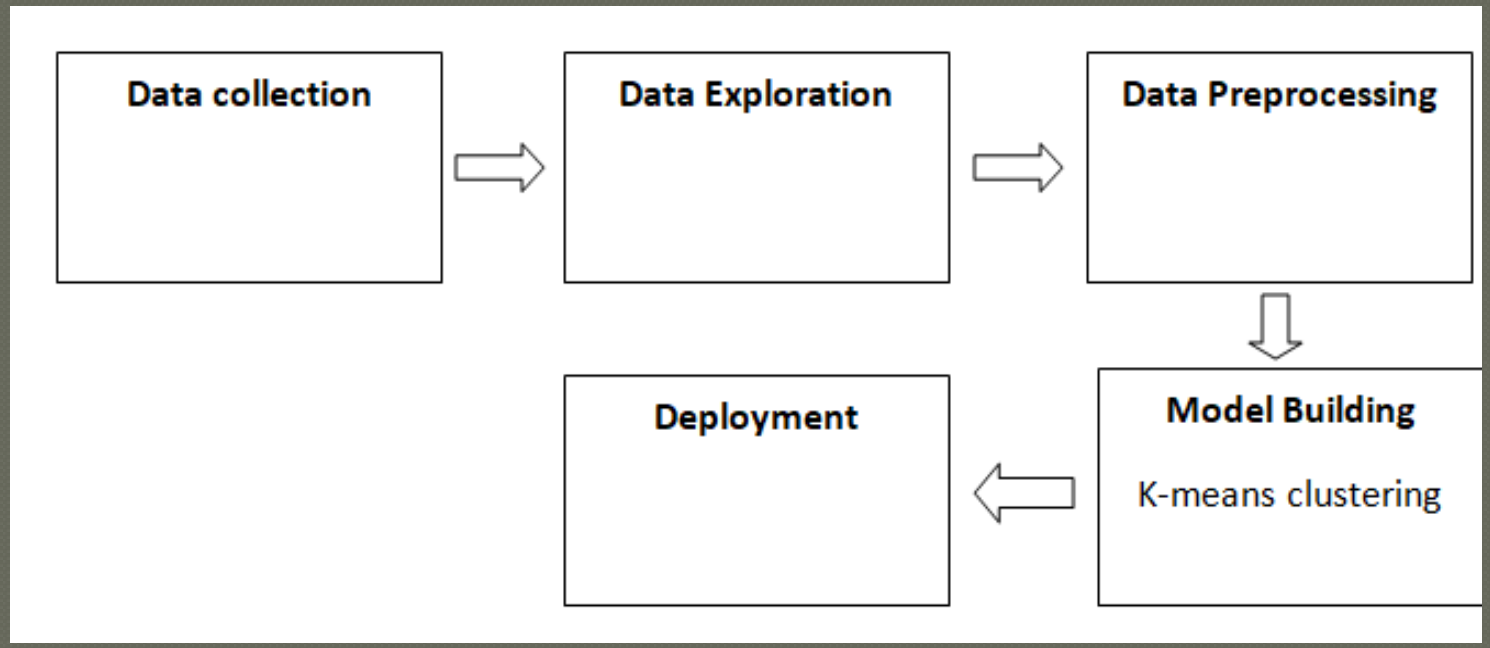
Objective:

- To create a solution to find groups of mall customers based on their spending behavior and annual income

Benefits:

- Can target specific groups of shoppers
- Can apply marketing strategies to make them shop more

Methodology



Model Training:

- Dataset

  The dataset for the task is downloaded from Kaggle in csv format for model training

- Data Preprocessing

  - Performing data exploration to get insight of data like understanding trends in the data etc.
  - Replaced outliers with the upper limit
  - Performed standardization on variables

Clustering –

- After data preprocessing, the data is fed into the clustering model for prediction.

- Elbow method and silhouette score method were used to find the ideal number of clusters

- The ideal number of clusters were then applied to the K-means clustering model to find groups

Clustering

- It was found that 5 was the apt number of clusters depending upon the elbow method and silhouette score
- The customers were grouped into 5 different groups based on their spending score and annual income.

Q & A:

Q1) What’s the source of data?

The data for training can be obtained from Kaggle in csv format

Q 2) What was the type of data?

The dataset was a combination of numerical and Categorical variables.

Q 3) What’s the complete flow you followed in this Project?

The project started with data collection, then data exploration, data preprocessing, model building, and finally deployment

Q 4) What techniques were you using for data pre-processing?

- Visualizing relation between the dependent and independent variables
- Replacing outlier values.
- Standardization of varaiables
- Removing redundant columns

Q 5) How training was done or what models were used?

- K-means clustering was used as the primary method for clustering and the ideal number of clusters were found using elbow method and silhouette score.

Q 6) What are the different stages of deployment?

The model was deployed as an API using FastAPI where the user can input values to the relevant variables to find which group the customer belongs to.