National Institute of Technology Karnataka, Surathkal
Indian Society For Technical Education
NITK Chapter

ML SGP Assignment (Week-3)

Q1) To get hands on experience and understanding of K-Means algorithm, Perform Clustering to the following dataset:

Dataset: Simple_Dataset
Dataset has coordinates of 20 points (x,y)

You may refer to the playlist or the following notebook which was used to explain in the playlist.
Notebook: Colab Notebook ML-Clustering

Q2) Perform K-Means Clustering for the following dataset, also known as Customer segmentation problems, (Take Number of clusters = 5)

Curious why k = 5 is ideal? Stay tuned.. We will discuss how to find the ideal number of clusters in the interactive session.

Dataset: Dataset.csv

Dataset has 5 columns:
Customer_ID: Unique ID assigned to the customer
Gender: Gender of the customer
Age: Age of the customer
Annual Income: Annual Income of the customer
Spending Score (1-100): Score assigned by the mall based on customer behavior and spending nature

(Hint: Don't forget to drop unnecessary columns before Clustering)

X axis: Annual Income, Y axis: SpendingScore, Z axis: Age
To visualize the Clustering, You may use Axes3D
(`from mpl_toolkits.mplot3d import Axes3D`)