AI & ML Internship - Task 8: K-Means Clustering

Objective

Perform unsupervised learning using K-Means clustering on customer segmentation data.

Dataset Used

- Mall_Customers.csv (mock customer segmentation dataset)
- Features used: 'Annual Income (k$)', 'Spending Score (1-100)'

Steps Performed

1. Loaded and explored the dataset.
2. Selected and scaled relevant features using StandardScaler.
3. Applied the Elbow Method to determine the optimal number of clusters (K).
4. Trained a KMeans model with the optimal K.
5. Assigned cluster labels to each data point.
6. Visualized the clusters using a scatter plot.
7. Evaluated clustering quality using the Silhouette Score.
8. Used PCA to reduce dimensions for visualizing cluster boundaries in 2D.

Tools & Libraries Used

- Python
- Pandas
- Matplotlib
- Seaborn
- Scikit-learn

Files Included

- Task8_KMeans_Clustering.ipynb (Colab notebook)

- Mall_Customers.csv (dataset)

- README.md (this document)

Status

Task Completed and Submitted.