Assignment

K-Means and DBSCAN

Objective:
The objective of this assignment is to explore and compare two popular clustering algorithms, K-Means and DBSCAN, and gain hands-on experience with their implementation and application.

Tasks:

1. Read about K-Means and DBSCAN algorithms and understand their working principles and differences.
2. Generate a synthetic dataset with 3-4 clusters using Scikit-Learn's make_blobs function.
3. Implement K-Means clustering algorithm on the generated dataset with varying values of K, i.e., 2, 3, 4, and visualize the resulting clusters using Matplotlib.
4. Evaluate the K-Means model using Silhouette score and determine the optimal number of clusters.
5. Implement DBSCAN clustering algorithm on the same dataset and visualize the resulting clusters.
6. Experiment with different values of hyperparameters, i.e., eps and min_samples, and observe their effect on the resulting clusters. 7. Compare the performance and limitations of K-Means and DBSCAN algorithms and provide your insights.

Deliverables:

1. Python code implementing K-Means and DBSCAN clustering algorithms and generating visualizations.

Note: You can use any programming language and tools of your choice, but Python, Matplotlib, and Scikit-Learn are recommended. You can also use real-world datasets or generate your own datasets for experimentation.

Data set
Link-https://drive.google.com/file/d/1XkoLTkJMy2pheCzkDEio03FrWaYvHw5l/view?usp=sharing