THEORETICAL ASSIGNMENT
NAME : MEETRAJSINH CHUDASAMA
MODULE 7 : UNSUPERVISED MACHINE LEARNING

Q-1. Explain Clustering Algorithms and Their Use Cases.

Clustering algorithms group similar data points together based on their characteristics. The goal is to identify groups, or clusters, of data points that are similar to each other, while being distinct from other groups. Some popular clustering algorithms include K-Means, Hierarchical clustering, and DBSCAN.

1. K-Means Clustering Algorithm :

The K-Means clustering algorithm is an iterative process where you are trying to minimize the distance of the data point from the average data point in the cluster.

2. Hierarchical Clustering Algorithms :

Hierarchical clustering algorithms seek to create a hierarchy of clustered data points. The algorithm aims to minimize the number of clusters by merging those closest to one another using a distance measurement such as Euclidean distance for numeric clusters or Hamming distance for text.

3. DBSCAN (Density-Based Spatial Clustering of Applications with Noise) :

DBSCAN identifies discrete groups in data. The algorithm aims to cluster the data as contiguous regions having high point density. Each cluster is separated from the others by points of low density. In simpler words, the cluster covers the data points that fit the density criteria which is the minimum number of data objects in a given radius.

Use Cases :

1. Market segmentation :

Businesses need to segment their market into smaller groups to understand the target audience. Clustering groups the like-minded people considering the neighborhood to generate similar recommendations, and it helps in pattern building and insight development.

2. **Retail marketing and sales :**

Marketing utilizes clustering to understand the customers' purchase behavior to regulate the supply chain and recommendations. It groups people with similar traits and probability of purchase. It helps in reaching the appropriate customer segments and provides effective promotions

3. **Social network analysis :**

Examining qualitative and quantitative social arrangements using network and Graph Theory. Clustering is required to observe the interaction amongst participants to acquire insights regarding various roles and groupings in the network.

4. **Wireless network analysis or Network traffic classification :**

Clustering groups together characteristics of the network traffic sources. Clusters are formed to classify the traffic types. Having precise information about traffic sources helps to grow the site traffic and plan capacity effectively.

5. **Image compression :**

Clustering helps store the images in a compressed form by reducing the image size without any quality compromise.

Q-2. Model Evaluation Techniques in Unsupervised Learning.

Clustering algorithms are pivotal tools in data analysis, enabling the identification of natural groupings within datasets. Key algorithms such as K-means, hierarchical, and DBSCAN offer distinct approaches to clustering. K-means, a widely-used centroid-based algorithm, partition data into K clusters by iteratively optimizing cluster centroids to minimize intra-cluster variance. Hierarchical clustering, conversely, constructs a tree-like hierarchy of clusters, either agglomeratively or divisively, allowing for the exploration of nested clusters. Meanwhile, DBSCAN, a density-based algorithm, groups together data points based on their proximity and density, distinguishing between core points, border points, and noise.