ML ASSIGNMENT ANSWERS

QUESTIONS	ANSWERS
1	D
2	D
3	C
4	B
5	D
6	C
7	D
8	A
9	A
10	B
11	A
12	B

Q13. What is the importance of clustering?

1. Clustering helps in understanding the natural grouping in a dataset. Their purpose is to make sense to partition the data into some group of logical groupings.

2. Its quality depends upon the methods and the identification of hidden patterns.

3. It helps to determine the internal structure of the data and restarting the local search procedure & removing the inefficiency.

4. This clustering analysis has been used for model analysis, vector region of attraction.

5. They play a wide role in applications like marketing economic research and weblogs to identify similarity measures, Image processing, and spatial research.

6. They are used in outlier detections to detect credit card fraudulence.

Q14. How can I improve my clustering performance?

The design of the model can be changed by tuning the hyperparameters. For K- Means clustering there are 3 main hyperparameters to set-up to define the best configuration of the model:

- Initial values of clusters
- Distance measures
- Number of clusters

Initial values of clusters greatly impact the clustering model, there are various algorithms to initialize the values. Distance measures are used to find points in clusters to the cluster centre, different distance measures yield different clusters. The number of clusters (k) is the most important hyperparameters in K-Means clustering. If we already know beforehand, the number of clusters to group the data into, then there is no use to tune the value of k.

For example,

k=10 for the MNIST digit classification dataset. If there aa is no idea about the optimal value of k, then there are various methods to find the optimal/best value of k.