MACHINE LEARNING Assisgnment 3

1. Which of the following is an application of clustering?
Ans:Option D: All of the above

2. On which data type, we cannot perform cluster analysis?
Ans: Option D: None

3. Netflix's movie recommendation system uses
Ans: Option C: Reinforcement learning and Unsupervised learning

4. The final output of Hierarchical clustering is
Ans: Option B: The tree representing how close the data points are to each other

5. Which of the step is not required for K-means clustering?
Ans: Option D. None

6. Which is the following is wrong?
Ans: Option C: k-nearest neighbour is same as k-means

7. Which of the following metrics, do we have for finding dissimilarity between two clusters in hierarchical clustering?
Ans:Option D: 1,2 and 3

8. Which of the following are true?
Ans: Option D: None of them

9)In the figure above, if you draw a horizontal line on y-axis for y=2. What will be the number of clusters formed?
Ans: Option A: 2

10. For which of the following tasks might clustering be a suitable approach?
Ans: Option C:. Predicting whether stock price of a company will increase tomorrow.

11. Given, six points with the following attributes:Which of the following clustering representations and dendrogram depicts the use of MIN or Single link proximity function in hierarchical clustering:
Ans: Option A

12. Given, six points with the following attributes:Which of the following clustering representations and dendrogram depicts the use of MAX or Complete link proximity function in hierarchical clustering:
Ans: Option B

13. What is the importance of clustering?
Ans: Clustering is a type of unsupervised machine learninng algorithm which helps to predict data with the help on only features provided. Due to which it allows us to find relationships between the data points based on the common attributes in the data and hence gives the output.

14. How can I improve my clustering performance?
Ans: Since we need to specify centroid in Kmeans clustring, hence we can improve it by changing the values of centroids and re-starting the analysis which will eventually increase the silhouette score.