1. Which of the following is an application of clustering?

d. All of the above


2. On which data type, we cannot perform cluster analysis?

d. None


3. Netflix's movie recommendation system uses

c. Reinforcement learning and Unsupervised learning


4. The final output of Hierarchical clustering is

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?

d. None

6. Which is the following is wrong

 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? i. Single-link ii. Complete-link iii. Average-link

d. 1, 2 and 3

8. Which of the following are true? i. Clustering analysis is negatively affected by multicollinearity of features ii. Clustering analysis is negatively affected by heteroscedasticity

 b. 2 only

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?

a. 2

10. For which of the following tasks might clustering be a suitable approach?

a.Given sales data from a large number of products in a supermarket, estimate future sales for each of these products.

11. Given, six points with the following attributes:

a.

12. Given, six points with the following attributes:

a.

13. What is the importance of clustering?

Clustering is important in data analysis and data mining applications. It is the task of grouping a set of objects so that objects in the same group are more similar to each other than to those in other groups (clusters).

14. How can I improve my clustering performance?

Graph-based clustering performance can easily be improved by applying ICA blind source separation during the graph Laplacian embedding step. Applying unsupervised feature learning to input data using either RICA or SFT, improves clustering performance.