

MACHINE LEARNING

1. Which of the following is an application of clustering?
Ans: d. All of the above
2. On which data type, we cannot perform cluster analysis?
Ans: d. None
3. Netflix's movie recommendation system uses –
Ans: a. Supervised learning
4. The final output of Hierarchical clustering is-
Ans: 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: d. None
6. Which of the following is wrong?
Ans: 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-linkAns: 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 heteroscedasticityAns: a. 1 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?
Ans: 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.
 - b. Given a database of information about your users, automatically group them into different market segments.
 - c. Predicting whether stock price of a company will increase tomorrow.
 - d. Given historical weather records, predict if tomorrow's weather will be sunny or rainy.

Ans: 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:

Ans: a

12. Given, six points with the following attributes:

Ans: b

13. What is the importance of clustering?

Ans: The primary use of clustering in machine learning is to extract valuable inferences from many unstructured data sets. If you are working with large amounts of data that are also not structured, it is only logical to organize that data to make it helpful in so many other ways, and clustering helps us do that.

Clustering is a significant component of machine learning, and its importance is highly significant in providing better machine learning techniques.

14. How can I improve my clustering performance?

Ans: Applying unsupervised feature learning to input data using either RICA or SFT, improves clustering performance. High clustering performance can be achieved by simply performing k-means clustering on the ICA components after PCA dimension reduction on the input data.