

# **MACHINE LEARNING**

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

- a. Biological network analysis
- b. Market trend prediction
- c. Topic modeling
- d. All of the above

**Ans :- b. Market trend prediction**

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

- a. Time series data
- b. Text data
- c. Multimedia data
- d. None

**Ans :- d. None**

**3. Netflix's movie recommendation system uses-**

- a. Supervised learning
- b. Unsupervised learning
- c. Reinforcement learning and Unsupervised learning
- d. All of the above

**Ans :- d. All of the above**

**4. The final output of Hierarchical clustering is-**

- a. The number of cluster centroids
- b. The tree representing how close the data points are to each other
- c. A map defining the similar data points into individual groups
- d. All of the above

**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?**

- a. A distance metric
- b. Initial number of clusters
- c. Initial guess as to cluster centroids
- d. None

**Ans :- d. None**

**6. Which of the following is wrong?**

- a. k-means clustering is a vector quantization method
- b. k-means clustering tries to group n observations into k clusters
- c. k-nearest neighbour is same as k-means
- d. None

**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-link

**Ans :- d. 1, 2 and 3**

**8. Which of the following are true?**

**Ans :- 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?**

**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 :- Option a.**

**12. Given, six points with the following attributes:**

**Ans :- Option b.**

**13. What is the importance of clustering?**

**Ans :-** 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). Clustering can be done by the different number.

**14. How can I improve my clustering performance?**

**Ans :-** There are several methods to improve the performance clustering algorithm First of all try to compare with one that works well and compare the results. Secondly,compare the time between the both algorithms,if u have set of good answers,then you can analyze how quality of the solution improves with the time Thirdly,try your algorithms with several instances of the problem one not too challenging ,one medium and one hard and finally, using evolution to optimize clustering parameters of your clustering algorithms.