MACHINE LEARNING ASSIGNMENT

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

ANS.B

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

ANS. D

3. Netflix's movie recommendation system uses?

ANS. D

4. The final output of Hierarchical clustering is?

ANS. B

5. Which of the step is not required for K-means clustering?

ANS. D

6. Which is the following is wrong?

ANS. C

7. Which of the following metrics, do we have for finding dissimilarity between two clusters in hierarchical clustering?

ANS. D

8. Which of the following are true?

ANS. B

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

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

ANS. A

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: 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. A good clustering algorithm is able to identity clusters irrespective of their shapes.

14. How can I improve my clustering performance?

ANS: There are several method to effectively assess the performance of your clustering algorithm.

First of all try to compare it against once that is known to work well. Then compare the results. Secondly, time your algorithms and compare the time between both algorithms. If you have two set of good answers, then you can analyse how the quality of the solution improves through time. Thirdly, try your algorithms with several instances of a problem. One not too challenging, one medium and one very hard. Finally, using the evolution to optimise the parameters of your clustering algorithms, could test it well under duress and could point to some direction of how to improve it.