

MACHINE LEARNING

Worksheet 3

1. d. All of the above
2. d. None
3. c. Reinforcement learning and Unsupervised learning
4. b. The tree representing how close the data points are to each other
5. d. None
6. c. k-nearest neighbour is same as k-means
7. d. 1, 2 and 3
8. a. 1 only
9. a. 2
10. b. .Given a database of information about your users, automatically group them into different market segments.
11. A.
12. B
13. a. Having clustering methods helps in restarting the local search procedure and remove the inefficiency. In addition, clustering helps to determine the internal structure of the data.
b. This clustering analysis has been used for model analysis, vector region of attraction.
c. Clustering helps in understanding the natural grouping in a dataset. Their purpose is to make sense to partition the data into some group of logical groupings.
d. Clustering quality depends on the methods and the identification of hidden patterns.
e. They play a wide role in applications like marketing economic research and weblogs to identify similarity measures, Image processing, and spatial research.
f. They are used in outlier detections to detect credit card fraudulence.
14. There are two important elements in improving the quality of clustering: improving the weights of the features in a document vector and creating a more appropriate distance measure. A good weighting technique can promote the good features of an object, and an appropriate distance measure can help bring similar features together.
 - a. A good document vector has the right kind of features, with higher weights assigned to the more important ones. In text data, there are two ways to improve the quality of a document vector: by removing noise and using a good weighting technique.