

## **Answers for Machine Learning Worksheet-3**

1. d. All of the above
2. d. None
3. d. All of the above
4. b. The tree representing how close the data points are to each other
5. d. None
6. c. k-nearest neighbor 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. The importance of clustering is as follows:
  - It helps us to understand the hidden structure within the data.
  - Density based clustering techniques can be applied to find out outliers and extreme points.
  - For particular data, if various clusters can be formed then for better strategy, clustering can be first applied before applying

supervised ML models to form a different model for all the clusters formed.

14. Clusters can be profiled with suitable graphs (with maximum 3 given input variables). It involves in understanding the clusters formed visually through graphs or analytically to draw inferences from the clusters with respect to the given input variables.

15. Clustering performances can be improved by following ways:

- Choosing appropriate input variables that can form groups more easily within the data.
- By choosing appropriate clustering algorithm for a given data.
- By tuning the hyperparameters for a given clustering algorithm and testing based on **Silhouette Coefficient**.