Machine Learning Assignment - 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. b. 2
10. b. Given a database of information about your users, automatically group them into different market segments.
11. a.
12. b.
13. Importance of Clustering :
    - 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.
    - 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.
    - Clustering quality depends on the methods and the identification of hidden patterns.
    - They play a wide role in applications like marketing economic research and weblogs to identify similarity measures, Image processing, and spatial research.

14. Ways to improve clustering performance :
    - Graph-based clustering performance can easily be improved by applying ICA blind source separation during the graph Laplacian embedding step.
    - Applying unsupervised feature learning to input data using either RICA or SFT, improves clustering performance.
    - For some cases, high clustering performance can be achieved by simply performing K-means clustering on the ICA components after PCA dimension reduction on the input data. However, the number of PCA and ICA signals/components needs to be limited to the number of unique classes.