1. d. All of the above
2. d. None
3. a. Supervised 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. 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. Clustering quality depends on the methods and the identification of hidden patterns.
14. The ways we can improve clustering performance are
15. Graph-based clustering performance can easily be improved by applying ICA blind source separation during the graph Laplacian embedding step.
16. Applying unsupervised feature learning to input data using either RICA or SFT, improves clustering performance.
17. Surprisingly 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.