# MACHINE LEARNING ASSIGNMENT-3

Ans1) All of the above (d)

Ans2) None (d)

Ans3) Reinforcement learning or unsupervised learning (c)

Ans4) The tree representing how close the data points are to each other (b)

Ans5) None(d)

Ans6) k-nearest neighbor is same as k-means (c)

Ans7) 1, 2 and 3 (d)

Ans8) 1 only (a)

Ans9) 2(a)

Ans10) Given a database of information about your users, automatically group them into different market segments (b)

Ans11) Option A

Ans12) Option B

Ans13) The importance of clustering is to make sense of and extract value from large sets of structured and unstructured data. If you're working with huge volumes of unstructured data, it only makes sense to try to partition the data into some sort of logical groupings before attempting to analyze it. Clustering and classification allow you to take a sweeping glance of your data en masse and then form some logical structures based on what you find there before going deeper into the nuts-and-bolts analysis.
In their simplest form, clusters are sets of data points that share similar attributes, and clustering algorithms are the methods that group these data points into different clusters based on their similarities. You'll see clustering algorithms used for disease classification in medical science, but you'll also see them used for customer classification in marketing research and for environmental health risk assessment in environmental engineering.

Ans14) K-means clustering algorithm can be significantly improved by

using a better initialization technique, and by repeating (re-starting) the algorithm.

When the data has overlapping clusters, k-means can improve the results of the initialization technique.

When the data has well-separated clusters, the performance of k-means depends completely on the goodness of the initialization.

Initialization using a simple furthest point heuristic (Maxmin) reduces the clustering error of k-means from 15% to 6%, on average.