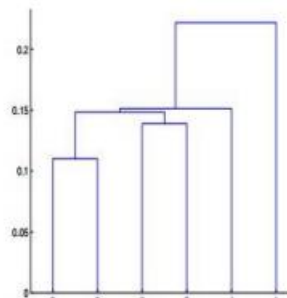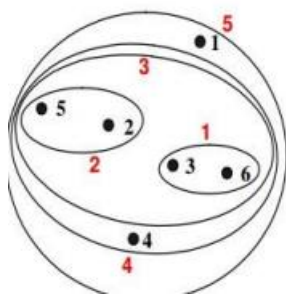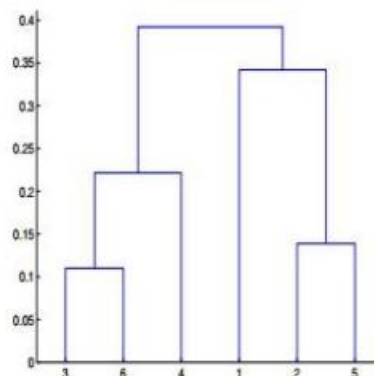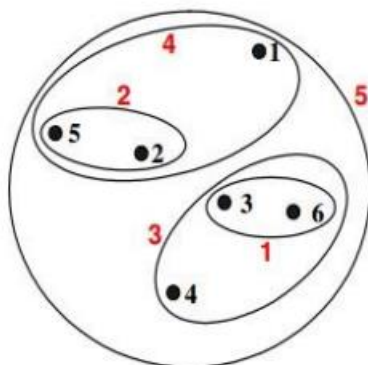# MACHINE LEARNING

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   a) Given sales data from a large number of products in a supermarket, estimate future sales for each of these products

11 a)
12 b)

13) Clustering is very much important as it determines the intrinsic grouping among the unlabelled data present. There are no criteria for good clustering. It depends on the user, what is the criteria they may use which satisfy their need. For instance, we could be interested in finding representatives for homogeneous groups (data reduction), in finding "natural clusters" and describe their unknown properties ("natural" data types), in finding useful and suitable groupings ("useful" data classes) or in finding unusual data objects (outlier detection). This algorithm

must make some assumptions that constitute the similarity of points and each assumption make different and equally valid clusters.

14) k-means suffers from its random initialization. Depending on which points get chosen as the start centers, the solution can be a very bad local minimum. And since k-means has a strictly concave loss function, it has no way of escaping from this local minimum during training.
A solution for this problem is the k-means++ algorithm, which uses a different initialization. The idea is pretty simple:
Instead of random initialization, we only choose the first center randomly. All following centers are then still sampled, but with a probability that is proportional to their squared distance from all current centers. Points furtheraway from current centers get a higher probability to become a center in the next iteration of initialization.
As we have seen, initialization can be key for the performance of k-means. Thek-means++ algorithm is a simple and widely applied technique to alleviate the problems that k-means has.