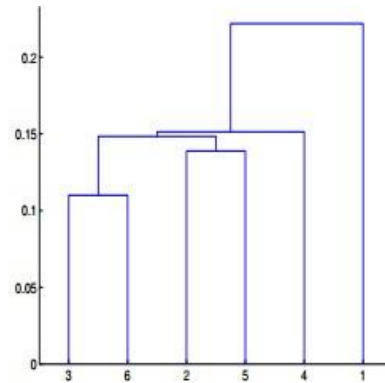
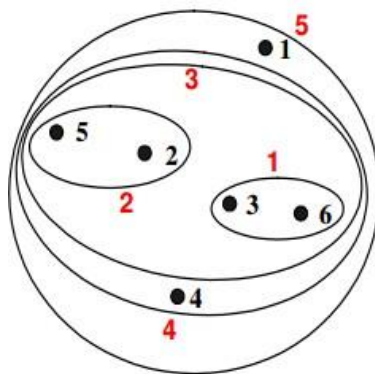


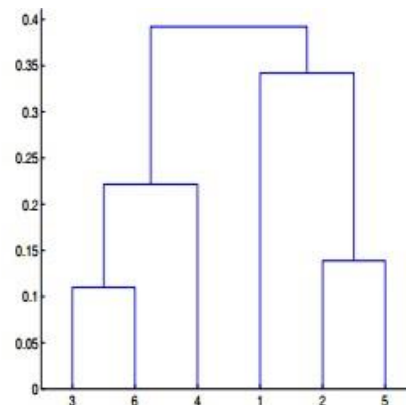
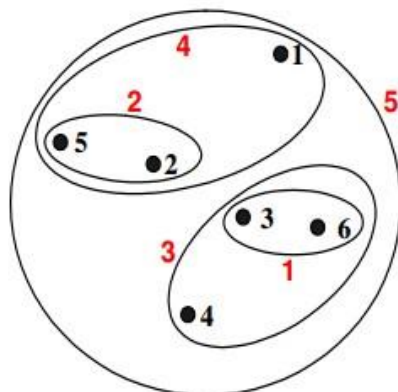
MACHINE LEARNING

Answer Keys

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 neighbor is same as k-means
7. d) 1,2 and 3
8. a) 1 only
9. a) 2
10. a) Assign each point to its nearest cluster
c) Update the cluster centroids based the current assignment

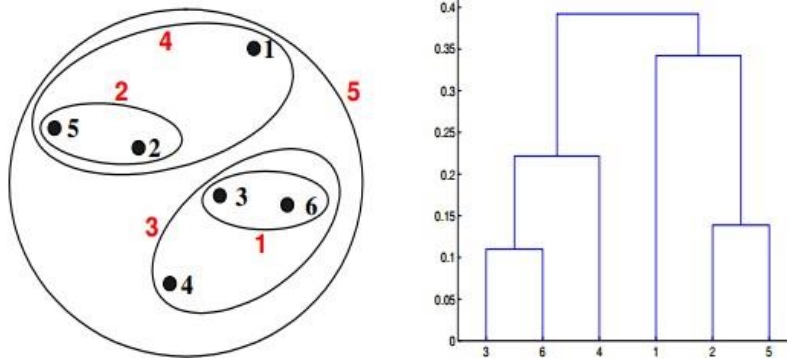


11. b) Explanation: For the single link or MIN version of hierarchical clustering, the proximity of two clusters is defined to be the minimum of the distance between any two points in the different clusters. For instance, from the table, we see that the distance between points 3 and 6 is 0.11, and that is the height at which they are joined into one cluster in the dendrogram. As another example, the distance between clusters $\{3, 6\}$ and $\{2, 5\}$ is given by $\text{dist}(\{3, 6\}, \{2, 5\}) = \min(\text{dist}(3, 2), \text{dist}(6, 2), \text{dist}(3, 5), \text{dist}(6, 5)) = \min(0.1483, 0.2540, 0.2843, 0.3921) = 0.1483$.



MACHINE LEARNING

12. b) Explanation: For the single link or MAX version of hierarchical clustering, the proximity of two clusters is defined to be the maximum of the distance between any two points in the different clusters. Similarly, here points 3 and 6 are merged first. However, $\{3, 6\}$ is merged with $\{4\}$, instead of $\{2, 5\}$. This is because the $\text{dist}(\{3, 6\}, \{4\}) = \max(\text{dist}(3, 4), \text{dist}(6, 4)) = \max(0.1513, 0.2216) = 0.2216$, which is smaller than $\text{dist}(\{3, 6\}, \{2, 5\}) = \max(\text{dist}(3, 2), \text{dist}(6, 2), \text{dist}(3, 5), \text{dist}(6, 5)) = \max(0.1483, 0.2540, 0.2843, 0.3921) = 0.3921$ and $\text{dist}(\{3, 6\}, \{1\}) = \max(\text{dist}(3, 1), \text{dist}(6, 1)) = \max(0.2218, 0.2347) = 0.2347$.



13. Clustering is important in data analysis and data mining applications. It is the task of grouping a set of objects so that objects in the same group are more similar to each other than to those in other groups (clusters). Clustering can be done by the different no.

14. 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.