MACHINE LEARNING ASSIGNMENT 2

Ans 1. d) 2 and 3

Ans 2. d) 1, 2 and 4

Ans 3. a) True

Ans 4. a) 1 only

Ans 5. b) 1

Ans 6. b) No

Ans 7. a) Yes

Ans 8. d) All of the above

Ans 9. a) K-means clustering algorithm

Ans 10. d) All of the above

Ans 11. d) All of the above

Ans 12. The K-means clustering algorithm is sensitive to outliers, because a mean is easily influenced by extreme values. K-medoids clustering is a variant of K-means that is more robust to noises and outliers. Instead of using the mean point as the center of a cluster, K-medoids uses an actual point in the cluster to represent it. Medoid is the most centrally located object of the cluster, with minimum sum of distances to other points.

Ans 13. Advantages of K means: -

- Relatively simple to implement.
- Scales to large data sets.
- Guarantees convergence.
- Can warm-start the positions of centroids.
- Easily adapts to new examples.
- Generalizes to clusters of different shapes and sizes, such as elliptical clusters.

Ans 14. The basic k-means clustering is based on a non-deterministic algorithm. This means that running the algorithm several times on the same data, could give different results. However, to ensure consistent results, k-means clustering using a deterministic method.