MACHINE LEARNING ASSIGNMENT

Answer 1. a) 2 Only

Answer 2. d) 1, 2 and 4

Answer 3 a) True

Answer 4 a) 1 only

Answer 5 b) 1

Answer 6 b) No

Answer 7 a) Yes

Answer 8 d) All of the above

Answer 9 a) K-means clustering algorithm

Answer 10 d) All of the above

Answer 11 d) All of the above

Subjective Answers

Answer 12. In K family the K- Means is sensitive to outliner. K-Means clustering algorithm is most sensitive to outliers as it uses the mean of cluster data points to find the cluster center and a mean is greatly influenced by outliner and thus cannot represent the correct cluster center.

Answer 13. K mean is better because:

- It is one of the simplest algorithm which uses unsupervised learning method to solve known clustering issues.
- It works really well with large datasets.

- If your data has no labels (class values or targets) or even column headers, K-Means will still successfully cluster your data.
- K Means clustering is found to work well when the structure of the clusters is hyper spherical (like circle in 2D, sphere in 3D).

Question 14. No, K means is a non-deterministic algorithm. This means that running the algorithm several times on the same data could give different results. We can proposed an improved, density-based version of K-means that includes a novel and systematic method of choosing initial centroids.