

Answer of Worksheet 2

1 – a

1 – d

3 – a

4 – a

5 – a

6 – b

7 – a

8 – d

9 – a

10 – d

11 – d

12 - **The k-means algorithm is sensitive to the outliers.** In this paper, we propose a robust two-stage k-means clustering algorithm based on the observation point mechanism, which can accurately discover the cluster centers without the disturbance of outliers.

13 - **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.

14 -The **non-deterministic** nature of K-Means is due to its random selection of data points as initial centroids. Method: We propose an improved, density based version of K-Means, which involves a novel and systematic method for selecting initial centroids.