|  |  |
| --- | --- |
|  | B |
|  | D |
|  | A |
|  | A |
|  | B |
|  | B |
|  | A |
|  | D |
|  | A |
|  | D |
|  | D |

Machine Learning Worksheet

Answer 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 centre 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

Answer 13

K-Means algorithm is **good in capturing structure of the data if clusters have a spherical-like shape**. It always try to construct a nice spherical shape around the centroid. K-Means is fast and scalable. Also in comparison to other clustering algorithms 1) If variables are huge, then K-Means most of the times computationally faster than hierarchical clustering, if we keep k smalls. 2) K-Means produce tighter clusters than hierarchical clustering, especially if the clusters are globular

Answer 14

No K-means is not deterministic, the non-deterministic nature of K-Means is due to its random selection of data points as initial centroids.