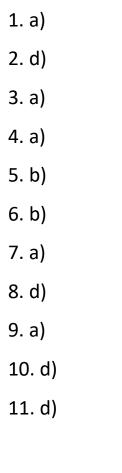
## Worksheet -2

## **Assignment-2 (Machine Learning)**



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.

## 13. K-means is better because: -

- 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.
- 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.

K-Means starts with a random set of data points as initial centroids. This random selection influences the quality of the resulting clusters.