

## ASSIGNMENT – 2

### MACHINE LEARNING

Q12 to Q14 are subjective answers type questions, Answers them in their own words briefly

Q12) Is K sensitive to outliers ?

Ans : 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.

Q13) Why is K means better?

Ans :

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

Q14) Is K means a deterministic algorithm?

Ans: This is non-deterministic nature of algorithms such as the K-Means clustering algorithm limits their applicability in areas such as cancer subtype prediction using gene expression data. It is hard to sensibly compare the results of such algorithms with those of other algorithms. The non-deterministic nature of K-Means is due to its random selection of data points as initial centroids