

## Worksheet 2

### Machine Learning

Answer -1

a) 2 only

Answer – 2

d) 1,2 and4

Answer – 3

a) true

Answer – 4

a) 1 only capping and flooring of variables

Answer – 5

b) 1

Answer – 6

c) No

Answer – 7

a) yes

Answer – 8

d) all of the above

Answer – 9

a) K-means clustering algorithms

Answer – 10

F ) All of the above

Answer – 11

D ) all of the above

Answer – 12

Yes , K-Means clustering algorithm is most sensitive to outliers as it uses the mean of cluster data points to find the cluster center, in other words because a mean is easily influenced by extreme values

Answer – 13

K-means clustering is better than others because

It is Relatively simple to implement.

It Scales to large data sets.

It Guarantees convergence.

It Can warm-start the positions of centroids.

It can Easily adapts to new examples.

It Generalizes to clusters of different shapes and sizes, such as elliptical clusters.

Answer-14

No, K- means is based on non-deterministic algorithm , it means that running the algorithm several times on the same data, could give different results.

The non-deterministic nature of K-Means is due to its random selection of data points as initial centroids.