**Assignment 2 machine learning**

1) a) 2 only

2) d) 1,2 and 4

3) a) true

4) a) Capping and flooring of variables

5) b) 1

6) a) yes

7) a) yes

8) d) all of the above

9) a) K-means clustering algorithm

10) d) All of the above

11) d) All of the above

**Brief**

12) YES, because the k-mean form the centre by taking the average of all the near values and when the outliers come into picture , it effects to the average.

For example: let’s say we have mean of 2,2,3,3,4,4,3,3,4,4,2,2 is 3

If we add a value 32 in it then the mean became 5.23 which is higher than all.

Hence, k means sensitive to outliers.

13) k-mean is one of the simplest algorithms which uses unsupervised learning method to solve known clustering issues. It works really well with large datasets.

14)Due to random selection of data points the running algorithm several times on the same data, could give you different results. Hence, K means is a non-deterministic algorithm.