MACHINE LEARNING WORKSHEET-2

Q1 b) 1 and 2

Q2 d) 1, 2 and 4

Q3 a) True

Q4 a) 1 only

Q5 b) 1

Q6 b) No

Q7 a) Yes

Q8 d) All of the above

Q9 a) K-means clustering algorithm

Q10 d) All of the above

Q11 d) All of the above

Q12 Yes, K-means is sensitive to outliers as when forming clusters with outliers present the centroid for each cluster tend to affect significantly. Also, the clusters formed with outliers present would be totally different than the dataset without outliers.

Q13 K-means is better because of the following reasons:

- It is simple to understand
- It is fast and easy to implement
- It can be used widely and can be scaled for large datasets

Q14 No. K-means is actually a non-deterministic algorithm because:

- It gives different result every time.
- Problem can not be solved in a given polynomial time.