Machine Learning Assignment-2
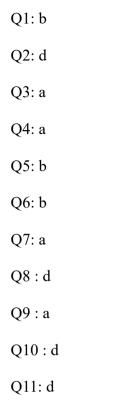
Q1: b

Q2: d

Q3: a

Q4: a

Q5: b

Q6: b

Q7: a

Q8 : d

Q9 : a

Q10 : d

Q11: d

Q12 : Yes, The K means clustering algorithm is sensitive to outliers because mean is easily influenced by extreme values and this is more robust to noises and outlier. K medoids uses an actual point in the cluster to represent it. In K means using the symmetric distance measure is the key component to define the samples that belonging to the same cluster.

Q13: It is really easy to implement and even more important most of the time you don't even have to implement it yourself and best unsupervised learning algorithm. This is one of the most widely used techniques for market or customer segmentation wherein the company data can be segregated into clusters and used to indentify certains patterns which leads to a more customised approach. Advantages of k means relatively simple to implement. It can warm start the positions of centroids and easily adapt to new examples.

Q14: 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. This algorithms usually have 2 steps 1) guessing steps 2) Assignemtn step. The K means algorithm divides the data space into k clusters such that the total variance of all data points with respect to the cluster mean is minimized.

WORKSHEET 2 SQL

Q1: D

Q2: A

Q3: C

Q4: D

Q5: B

Q6: D

Q7: A

Q8: C

Q9: D

Q10: D

Q11: A

Q12: C

Q13: A

Q14: B C D

Q15: A B

STATISTICS WORKSHEET -2

Q1: B

Q2: C

Q3: D

Q4: C

Q5: B

Q6: B

Q7: A

Q8: B

Q9: D

Q10: A

Q11: C

Q12: B

Q13: D

Q14: A

Q15: D