Assignment no 2

Q.1) b

Q.2) d

Q.3) b

Q.4) a

Q.5) b

Q.6) b

Q.7) a

Q.8) d

Q.9) a

Q.10) d

Q.11) d

Q.12) K Means is most sensitive to outliers because in this type we calculate mean of data and even one uneven entry can distort the mean. Since it is a distance based algorithm outliers can affect a lot.

Q.13) Advantages of K Means-

a) It is very easy to implement.

b) It can be used for huge data sets and works faster for large data sets.

c) It adapts the new examples very quickly.

d) It forms generalized clusters for different shapes and data sets

Q.14) The basic k-means clustering is based on a non-deterministic algorithm. This means that running the algorithm several times on the same data set, could give different results.