MACHINE LEARNING SET 2

A1) Option a- 2 Only

A2) Option d- 1, 2 and 4

A3) Option a- True

A4) Option a- 1 only

A5) Option b-1 only

A6) Option b- No

A7) Option a- Yes

A8) Option d- All of the above

A9) Option a- K-means clustering algorithm

A10) Option d- All of the above

A11) Option d- All of the above

A12) The K-means clustering algorithm is sensitive to outliers, because in k-means for assigning any observation to any cluster the distance is calculated between that observation and means of other clusters and then further a new mean is calculated so if any outlier observation is added to any cluster then it would affect the newly calculated mean thus affecting the cluster creation.

A13) K-means is better because of following reasons:

- Relatively simple to implement.

- Scales to large data sets.

- Guarantees convergence.

- Can warm-start the positions of centroids.

- Easily adapts to new examples.

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

A14) The k-means algorithm is non-deterministic in nature. The non-deterministic nature of K-Means is due to its random selection of data points as initial centroids.