MACHINE LEARNING ANSWER SHEET

Q-1:   A- 2 only

Q-2:   D- 1, 2 and 4

Q-3:   A- True

Q-4:   A- 1 Only

Q-5:   A- 0

Q-6:   B- No

Q-7:   A- Yes

Q-8:   D- All of the above

Q-9:   A- K-mean clustering algorithm

Q-10:   A- 1 Only

Q-11:   D- All of the above

Q-12:

The number of clusters found from data by the method is denoted by the letter 'K' in K-means. In this method, data points are assigned to clusters in such a way that the sum of the squared distances between the data points and the centroid is as small as possible.

The K-means clustering algorithm is sensitive to outliers, because a mean is easily influenced by extreme values. K-medoids clustering is a variant of K-means that is more robust to noises and outliers.

Q-13:

K-means has been around since the 1970s and fares better than other clustering algorithms like density-based, expectation-maximisation. It is one of the most robust methods, especially for image segmentation and image annotation projects. According to some users, K-means is very simple and easy to implement.

K means is relatively simple to implement, scales large data sets, guarantees convergence, easily adapts to new examples, generalizes to clusters of different shapes and sizes, such as elliptical clusters.

Q-14:

This is non-deterministic nature of algorithms such as the K-Means clustering algorithm limits their applicability in areas such as cancer subtype prediction using gene expression data. It is hard to sensibly compare the results of such algorithms with those of other algorithms.