**Machine Learning Worksheet - 2**

**Q.1** (b)

**Q.2** (d)

**Q.3** (a)

**Q.4** (a)

**Q.5** (b)

**Q.6** (b)

**Q.7** (a)

**Q.8** (d)

**Q.9** (d)

**Q.10** (a)

**Q.11** (d)

**Q.12** (d)

**Q.13** The K-means clustering algorithm is sensitive to outliers, because a mean is easily influenced by extreme values.

**Q.14** Advantages of k-means

1. Relatively simple to implement.

2. Scales to large data sets.

3. Guarantees convergence.

4. Can warm-start the positions of centroids.

5. Easily adapts to new examples.

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

**Q.15** 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. However, to ensure consistent results, FCS Express performs k-means clustering using a deterministic method.