**MACHINE LEARNING 1-WORKSHEET**

1. d
2. d
3. d
4. a
5. b
6. d
7. a
8. b
9. d
10. a
11. d
12. a
13. **How is cluster analysis calculated?**

* In the Cluster analysis  the task of grouping a set of objects or the parameters in the dataset in such a way that objects or the parameters in the same group are more similar to each other than to those in other groups, which is actually forms different cluster. The cluster analysis follows three basic steps: 1) calculate the distances between two points in the cluster, 2) link the clusters, and 3) Find the solution by selecting the right number of clusters. But in the beginning, we need to select the variables for which we formed the clusters.

1. **How is cluster quality measured?**

* I would see the quality of a clustering by generating data from known dataset generating processes and then check and find out how often patterns are misclassified by the clustering. But this involved making assumptions about the distribution of patterns from each generating steps, but we can use datasets designed for supervised classification. The quality of a clustering algorithm is also performed by its ability to find out some or all of the hidden patterns inside the dataset.

1. **What is cluster analysis and its types?**

**Definition:**

1. It groups the similar data/input variables in same group.
2. The objective of this procedure is that the objects in a group are similar to one another and are different from the objects in other groups that is closer to the mean value.
3. Higher the similarity within a group and greater difference between the groups, more distinct the clustering.
4. Cluster analysis gives a potential relationship and construct systematic structure in large number of variable’s and observations.

**Types:**

1. **Hierarchical clustering:**  it clusters to exist within larger clusters to form a tree.
2. **Partition clustering:**It’s simply a division of the set of data sets into non-overlapping clusters such that each objects is in exactly one subset.
3. **Exclusive Clustering:** They assign each value to a single cluster.
4. **Overlapping Clustering:** It is commonly reflect the fact that an object can simultaneously belong to more than one group in the overall dataset.
5. **Fuzzy clustering:**Every objects belongs to every cluster with a membership weight that goes between 0: if it absolutely doesn't belong to cluster and 1: if it absolutely belongs to the cluster.
6. **Complete clustering:** It perform a hierarchical clustering using a set of dissimilarities on 'n' objects that are being clustered. They tend to find compact clusters of an approximately equal diameter.