WORKSHEET-3

MACHINE_LEARNING_ANSWERSHEET_3

1. (d) All of the Above
2. (d) None
3. (c) Reinforcement learning and Unsupervised Learning
4. (b) The Tree representing how close the data points are to each other
5. (d) None
6. (c) K-nearest neighbor is same as K-Means
7. (d) 1,2 and 3
8. (a) Clustering Analysis is negatively affected by Multicollinearity.
9. (a) 2

11. (b)
12. (b)

13. IMPORATANCE OF CLUSTERING:

- **The primary use of clustering is to extract valuable inferences from many unstructured data sets.**
- **Clustering helps to determine the internal structure of the data.**
- **Clustering helps in understanding the natural grouping in a dataset. Their purpose is to make sense to partition the data into some group of logical grouping.**
- **Clustering quality depends on the methods and the identification of hidden patterns.**
- **Clustering is used in outlier detections to detect credit card fraudulence.**
- **Clustering plays a vital role in applications like marketing economic research and weblogs to identify similarity measures and spatial research.**

14. Using either RICA (Reconstruction independent component analysis) or SFT, we can apply unsupervised feature learning to input data to improve clustering performance. High Clustering performance can be achieved by simply performing K-means clustering on the ICA components after principal component analysis dimension reduction on the input data.