**NAME OF THE PROJECT**

**Machine learning  
Submitted by:**

**SAUNAK MUKHERJEE**



1. **Which of the following is an application of clustering?**

d. All of the above

1. **On which data type, we cannot perform cluster analysis?**

d. None

1. **Netflix’s movie recommendation system uses**

c. Reinforcement learning and Unsupervised learning.

**4. The final output of Hierarchical clustering is**

b. The tree representing how close the data points are to each other

**5. Which of the step is not required for K-means clustering?**

d. None

**6. Which is the following is wrong?**

c. k-nearest neighbour is same as k-means

**7. Which of the following metrics, do we have for finding dissimilarity between two clusters in**

**hierarchical clustering?**

i.

Single-link

ii.

Complete-link

iii.

Average-link

d. 1, 2 and 3

**8. Which of the following are true?**

**i. Clustering analysis is negatively affected by multicollinearity of features**

**ii. Clustering analysis is negatively affected by heteroscedasticity**

a. 1 only

**9. In the figure above, if you draw a horizontal line on y-axis for y=2. What will be the number of clusters**

**formed?**

a. 2

**10. For which of the following tasks might clustering be a suitable approach?**

a.Given sales data from a large number of products in a supermarket, estimate future sales for each

of these products.

**13. What is the importance of clustering?**

Clustering which can also **used for anomaly detection to find data points that are not part of any cluster, or outliers**. Clustering is used to identify groups of similar objects in datasets with two or more variable quantities.

**14. How can I improve my clustering performance?**

**We are Applying unsupervised feature learning to input data using either RICA or SFT**, improves clustering performance.