1. Among the following identify the one in which dimensionality reduction reduces.

(D) Collinearity

2. Which of the following machine learning algorithm is based upon the idea of bagging?

(B) Random Forest

3. Choose a disadvantage of decision trees among the following.

(C) Decision Tree are prone to overfit

4. What is the term known as on which the machine learning algorithms build a model based on sample data?

(C). Training data

5. Which of the following machine learning techniques helps in detecting the outliers in data?

c. Anomaly detection

6. Identify the incorrect numerical functions in the various function representation of machine learning.

c. Case based

7. Analysis of ML algorithm needs

d) Both a and b (Statistical learning theory and Computational learning theory)

8. Identify the difficulties with the k-nearest neighbor algorithm.

c. Both a and b (Curse of dimensionality and Calculate the distance of test case for all training cases)

9. The total types of the layer in radial basis function neural networks is_____

(B) 2

10. Which of the following is not a supervised learning

(a) PCA

11. What is unsupervised learning?

(c) Neither feature nor number of groups is known

12. Which of the following is not a machine learning algorithm?

(b) SVG

13. is the scenario when the model fails to decipher the underlying trend in the input data

(B) Underfitting

15. What is called the average squared difference between classifier predictions and actual values in regression problems?

(B) Mean squared error

16. Logistic regression is a..............regression technique that is used to model data having a .... outcome.

(A) Linear, binary

17. You are given reviews of few Netflix series marked as positive, negative and neutral. Classifying reviews of a new Netflix series is an example of

(a) Supervised learning

18. Following is powerful distance metrics used by Geometric model

(c) Both a and b (Euclidean distance and Manhattan distance)

19. Which of the following techniques would perform better for reducing dimensions of a data set?

(b) **Removing columns which have high variance in data**

20. Supervised learning and unsupervised clustering both require which is correct according to the statement.

(c) Input attribute.

21. What is the meaning of hard margin in SVM?

(a) SVM allows very low error in classification

15. Increase in which of the following hyperparameters results in overfit in Random forest?
    (1). Number of Trees. (2). Depth of Tree, (3). Learning Rate

(c) 2 and 3

16. Below are the 8 actual values of the target variable in the train file: [0,0,0, 0, 1, 1,1,1,1,1],
    What is the entropy of the target variable?

(B) 6/10 log(6/10) + 4/10 log(4/10)

25. Consider the problem of binary classification. Assume I trained a model on a linearly separable training set, and now I have a new labeled data point that the model properly categorized and is far away from the decision border. In which instances is the learnt decision boundary likely to change if I now add this additional point to my previous training set and re-train? When the training model is,

c) Either 1 or 3

26. Assume you've discovered multi-collinear features. Which of the following actions do you intend to take next? (1). Both collinear variables should be removed. (2). Instead of deleting both variables, we can simply delete one. (3). Removing correlated variables may result in information loss. We may utilize penalized regression models such as ridge or lasso regression to keep such variables.

B) increase by 5 pounds

27. A least squares regression study of weight (y) and height (x) yielded the following least squares line: y = 120 + 5x. This means that if the height is increased by one inch, the weight should increase by what amount?

d) Minimize the squared distance from the points

28. The line described by the linear regression equation (OLS) attempts to ____?

b) As the value of one attribute increases the value of the second attribute also increases

29. For two real-valued attributes, the correlation coefficient is 0.85. What does this value indicate?

b) Convolutional Neural Network