

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

- a) Performance
- b) Statistics
- c) Entropy
- d) Collinearity

Ans : d) Collinearity

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

- a) Decision Tree
- b) Random Forest
- c) Classification
- d) SVM

Answer : b) Random Forest

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

- a) Decision tree robust to outliers
- b) Factor analysis
- c) Decision Tree are prone to overfit
- d) all of the above

Answer : 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?

- a) Data Training
- b) Sample Data
- c) Training data
- d) None of the above

Answer : c) Training Data

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

- a) Clustering
- b) Classification
- c) Anamoly detection
- d) All of the above

Answer : c) Anamoly detection

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

- a) Support Vector
- b) Regression
- c) Case based
- d) Classification

Answer : c) Case based

7. Analysis of ML algorithm needs :

- a) Statistical learning theory
- b) Computational learning theory
- c) None of the above
- d) Both a and b

Answer : d) Both a and b

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

- a) Curse of dimensionality
- b) Calculate the distance of test case for all training cases
- c) Both a and b
- d) None of the above

Answer : c) Both a and b

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

- a) 1
- b) 2
- c) 3
- d) 4

Answer : c) 3

10. Which of the following is not a supervised learning

- a) PCA
- b) Naïve bayes
- c) Linear regression
- d) KMeans

Answer : a) PCA

11) What is unsupervised learning?

- a) Number of groups may be known
- b) Features of groups explicitly stated
- c) Neither feature nor number of groups is known
- d) None of the above

Answer : c) Neither feature nor number of groups is known.

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

- a) SVM
- b) SVG
- c) Random Forest Algorithm
- d) None of the above

Answer : b) SVG

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

- a) Overfitting
- b) Underfitting
- c) Both a and b
- d) None of the above

Answer : b) Underfitting

14) Real-Time decisions, Game AI, Learning Tasks, Skill acquisition, and Robot Navigation are applications of

- a) Reinforcement learning
- b) Supervised learning
- c) Unsupervised Learning
- d) None of the above

Answer : a) Reinforcement learning

15) What is called the average squared difference between classifier predicted output and actual output?

- a) Mean relative error
- b) Mean squared error
- c) Mean absolute error
- d) Root mean squared error

Answer : b) Mean squared error

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

- a) Linear, binary
- b) Linear, numeric
- c) Nonlinear, binary
- d) Nonlinear, numeric

Answer : c) Nonlinear, 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
- b) unsupervised learning
- c) semisupervised learning
- d) reinforcement learning

Answer : a) supervised learning

18) Following is powerful distance metrics used by Geometric model

- a) euclidean distance
- b) manhattan distance
- c) both a and b
- d) square distance

Answer : c) both a and b

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

- a) removing columns which have too many missing values
- b) removing columns which have high variance in data
- c) removing columns with dissimilar data trends
- d) none of these

Answer : a) removing columns which have too many missing values

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

- A. output attribute.
- B. hidden attribute.
- C. input attribute.
- D. categorical attribute

Answer : B. hidden attribute.

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

- (A) SVM allows very low error in classification
- (B) SVM allows high amount of error in classification
- (C) Underfitting
- (D) SVM is highly flexible

Answer : (A) SVM allows very low error in classification

22) Increase in which of the following hyper parameter results into overfit in Random forest?

(1). Number of Trees. (2). Depth of Tree, (3). Learning Rate

- (A) Only 1
- (B) Only 2
- (C) 2 and 3
- (D) 1,2 and 3

Answer : (B) Only 2

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

- (A) $-(6/10 \log(6/10) + 4/10 \log(4/10))$
- (B) $6/10 \log(6/10) + 4/10 \log(4/10)$
- (C) $4/10 \log(6/10) + 6/10 \log(4/10)$
- (D) $6/10 \log(4/10) - 4/10 \log(6/10)$

Answer : A) $-(6/10 \log(6/10) + 4/10 \log(4/10))$

Explanation : The formula for calculating the entropy is:

Entropy = $-(p(0) \times \log(P(0)) + p(1) \times \log(P(1)))$.

Total no of outcomes = $4+6 = 10$

Probability of getting an outcome 1 $p(0) = 6/10$

Probability of getting an outcome 0 $p(1) = 4/10$

Entropy = $-((6/10 \log(6/10) + 4/10 \log(4/10))$

24) Lasso can be interpreted as least-squares linear regression where

- (A) weights are regularized with the l1 norm
- (B) weights are regularized with the l2 norm
- (C) the solution algorithm is simpler

Answer : (A) weights are regularized with the l1 norm

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,

- (A) Perceptron and logistic regression
- (B) Logistic regression and Gaussian discriminant analysis
- (C) Support vector machine
- (D) Perceptron

Answer : (B) Logistic regression and Gaussian discriminant analysis

Explanation :

- Predicting the output variable, which is discrete in two classes, is referred to as binary classification.
- The method of modeling the probability of a discrete result given an input variable is known as logistic regression.
- The most frequent logistic regression models include a binary result, which can accept two values such as true/false or yes/no.
- A learning algorithm based on a probabilistic assumption is Gaussian Discriminant Analysis.

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.

- (A) Only 1
- (B) Only 2
- (C) Either 1 or 3
- (D) Either 2 or 3

Answer : (D) Either 2 or 3

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?

- (A) increase by 1 pound
- (B) increase by 5 pound
- (C) increase by 125 pound
- (D) None of the above

Answer : (B) increase by 5 pound

Explanation : $y = 120 + 5x$

$$x = x + 1$$

$$y_1 = 120 + 5(x + 1)$$

$$y_1 = 120 + 5x + 5$$

$$y_1 = 125 + 5x$$

$$\text{Difference } (y - y_1) = 5 \text{ pounds}$$

Hence, y increase by 5 pounds.

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

- (A) Pass through as many points as possible.
- (B) Pass through as few points as possible
- (C) Minimize the number of points it touches
- (D) Minimize the squared distance from the points

Answer : (D) Minimize the squared distance from the points

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

- (A) The attributes are not linearly related
- (B) As the value of one attribute increases the value of the second attribute also increases
- (C) As the value of one attribute decreases the value of the second attribute increases
- (D) The attributes show a curvilinear relationship

Answer : (C) As the value of one attribute decreases the value of the second attribute increases

30) Which neural network architecture would be most suited to handle an image identification problem (recognizing a dog in a photo)?

- (A) Multi Layer Perceptron
- (B) Convolutional Neural Network
- (C) Recurrent Neural network
- (D) Perceptron

Answer : (B) Convolutional Neural Network

Explanation :The leading architecture used for image recognition and detection tasks is **Convolutional Neural Networks** (CNNs). Convolutional neural networks consist of several layers with small neuron collections, each of them perceiving small parts of an image.