

File 3

MCQ

QUESTION 1)

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

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

Correct answer is option (d) Collinearity

QUESTION 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

Correct answer is option (b) Random Forest

QUESTION 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

Correct answer is option (c) Decision Tree are prone to overfit

QUESTION 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

Correct answer is option (c) Training data

QUESTION 5)

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

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

Correct answer is option (c) Anomaly detection

QUESTION 6)

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

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

Correct answer is option (c) Case based

QUESTION 7)

Analysis of ML algorithm needs

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

Correct answer is option (d) Both a and b

QUESTION 8)

Identify the difficulties with the k-nearest neighbour algorithm.

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

Correct answer is option (c) Both a and b

QUESTION 9)

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

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

Correct answer is option (c) 3

QUESTION 10)

Which of the following is not a supervised learning

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

Correct answer is option (a) PCA

QUESTION 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

Correct answer is option (c) Neither feature nor number of groups is known

QUESTION 12)

Which of the following is not a machine learning algorithm?

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

Correct answer is option (b) SVG

QUESTION 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

Correct answer is option (b) Underfitting

QUESTION 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

Correct answer is option (a) Reinforcement learning

QUESTION 15)

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

55) What is called the average squared difference between classifier

- a) Mean relative error
- b) Mean squared error

- c) Mean absolute error
- d) Root mean squared error

Correct answer is option (b) Mean squared error

QUESTION 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

Correct answer is option (c) Nonlinear, binary

QUESTION 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. semi-supervised learning
- D. reinforcement learning

Correct answer is option (A) supervised learning

QUESTION 18)

Following is powerful distance metrics used by Geometric model

- A. Euclidean distance
- B. Manhattan distance
- C. both a and b
- D. square distance

Correct answer is option (c) Both a and b

QUESTION 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

Correct answer is option (A) removing columns which have too many missing values

QUESTION 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

Correct answer is option (C. input attribute.

QUESTION 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

Correct answer is option (A) SVM allows very low error in classification

QUESTION 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

Correct answer is option (B) Only 2

QUESTION 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) $6/10 \log(4/10) - 4/10 \log(6/10)$

Correct answer is option (A) $-(6/10 \log(6/10) + 4/10 \log(4/10))$

QUESTION 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

Correct answer is option (A) weights are regularized with the l1 norm

QUESTION 25)

Consider the problem of binary classification. Assume I trained a model on a linearly separable training set, and now I have a new labelled 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

Correct answer is option (B) Logistic regression and Gaussian discriminant analysis

QUESTION 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

Correct answer is option (D) Either 2 or 3

QUESTION 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 pounds
- (C) increase by 125 pounds
- (D) None of the above

Correct answer is option (B) increase by 5 pounds

QUESTION 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

Correct answer is option (D) Minimize the squared distance from the points

QUESTION 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

Correct answer is option (C) As the value of one attribute decreases the value of the second attribute increases

QUESTION 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

Correct answer is option (B) Convolutional Neural Network