Unit-III

1.A collection of one or more items is called as _____

(A)Itemset

(B)Support

(C)Confidence

(D)Support Count

Ans:A

2.Frequency of occurrence of an itemset is called as _____

(A)Support

(B)Confidence

(C)Support Count

(D)Rules

Ans:C

3.An itemset whose support is greater than or equal to a minimum support threshold is _____

(A)Itemset

(B)Frequent Itemset

(C)Infrequent items

(D)Threshold values

Ans:B

4.What does FP growth algorithm do?

(A)It mines all frequent patterns through pruning rules with lesser support

(B)It mines all frequent patterns through pruning rules with higher support

(C)It mines all frequent patterns by constructing a FP tree

(D)It mines all frequent patterns by constructing an itemsets

Ans:C

5.What techniques can be used to improve the efficiency of apriori algorithm?

(A)Hash-based techniques

(B)Transaction Increases

(C)Sampling

(D)Cleaning

Ans:A

6. Linear Regression is a supervised machine learning algorithm.

A) TRUE

B) FALSE

Ans:A

7.It is possible to design a Linear regression algorithm using a neural network?

A) TRUE

B) FALSE

Ans:A

8.Which of the following methods do we use to find the best fit line for data in Linear Regression?

A) Least Square Error

B) Maximum Likelihood

C) Logarithmic Loss

D) Both A and B

Ans:A

9. A local retailer has a database that stores 10,000 transactions of lastsummer. After analyzing the data,a data science team has identified thefollowing statistics:• {battery} appears in 6,000 transactions.• {sunscreen}appears in 5,000 transactions.• {sandals} appears in 4,000 transactions.•{bowls} appears in 2,000 transactions.• {battery, sunscreen} appears in1,500 transactions.• {battery, sandals} appears in 1,000 transactions.•{battery, bowls} appears in 250 transactions.• {battery, sunscreen, sandals}appears in 600 transactions. Q) What are the confidence values of{battery}->{ sunscreen} and {battery, sunscreen}->{ sandals} ?

a) 0.3 and 0.4
b) 0.25 and 0.4
c) 0.25 and 0.15
d) 0.6 and 0.4
Ans: b

10.Which of the following implies no relationship with respect to correlation?

a) Cor(X, Y) = 1

b) Cor(X, Y) = 0

c) Cor(X, Y) = 2

d) All of the mentioned

Ans:b

11. If Linear regression model perfectly first i.e., train error is zero, then
_____

a) Test error is also always zero

b) Test error is non zero

c) Couldn’t comment on Test error

d) Test error is equal to Train error

Ans:C

12.Which of the following metrics can be used for evaluating regression models?

i) R Squared

ii) Adjusted R Squared

iii) F Statistics

iv) RMSE / MSE / MAE

a) ii and iv

b) i and ii

c) ii, iii and iv

d) i, ii, iii and iv

Ans:d

13.How many coefficients do you need to estimate in a simple linear regression model (One independent variable)?

a) 1

b) 2

c) 3

d) 4

Ans:b

14.In a simple linear regression model (One independent variable), If we change the input variable by 1 unit. How much output variable will change?

a) by 1

b) no change

c) by intercept

d) by its slope

Ans:d

15.Function used for linear regression in R is _____

a) lm(formula, data)

b) lr(formula, data)

c) lrm(formula, data)

d) regression.linear(formula, data)

Ans:a

16.In syntax of linear model lm(formula,data,..), data refers to _____

a) Matrix

b) Vector

c) Array

d) List

Ans:b

17.In the mathematical Equation of Linear Regression $Y = \beta 1 + \beta 2X + \epsilon$, ($\beta 1$, $\beta 2$) refers to _____

a) (X-intercept, Slope)

b) (Slope, X-Intercept)

c) (Y-Intercept, Slope)

d) (slope, Y-Intercept)

Ans:c

18. _____ is an incredibly powerful tool for analyzing data.

a) Linear regression

b) Logistic regression

c) Gradient Descent

d) Greedy algorithms

Ans:a

19.The square of the correlation coefficient r 2 will always be positive and is called the _____

a) Regression

b) Coefficient of determination

c) KNN

d) Algorithm

Ans:b

20.Predicting y for a value of x that's outside the range of values we actually saw for x in the original data is called _____

a) Regression

b) Extrapolation

c) Intrapolation

d) Polation

Ans:b

21.What is predicting y for a value of x that is within the interval of points that we saw in the original data called?

a) Regression

b) Extrapolation

c) Intrapolation

d) Polation

Ans:c

22. _____ is a simple approach to supervised learning. It assumes that the dependence of Y on X1, X2, . . . Xp is linear.

a) Linear regression

b) Logistic regression

c) Gradient Descent

d) Greedy algorithms

Ans:a

23.Although it may seem overly simplistic, _____ is extremely useful both conceptually and practically.

a) Linear regression

b) Logistic regression

c) Gradient Descent

d) Greedy algorithms

Ans:a

24. _____ refers to a group of techniques for fitting and studying the straight-line relationship between two variables.

a) Linear regression

b) Logistic regression

c) Gradient Descent

d) Greedy algorithms

Ans:a

25. What do you mean by support(A)?
a. Total number of transactions containing A
b. Total Number of transactions not containing A
c. Number of transactions containing A / Total number of transactions
d. Number of transactions not containing A / Total number of transactions

Ans: c