Started on	Friday, 24 September 2021, 6:00 PM
State	Finished
Completed on	Friday, 24 September 2021, 7:14 PM
Time taken	1 hour 13 mins
Marks	43.00/49.00
Grade	**87.76** out of 100.00

Question **1**

Complete

Mark 1.00 out of 1.00

Which of the following statements is/are not TRUE with respect to the multi variate optimization

I - The gradient of a function at a point is parallel to the contours

II - Gradient points in the direction of greatest increase of the function

III - Negative gradients points in the direction of the greatest decrease of the function

IV - Hessian is a non-symmetric matrix

Select one:

- **a. III and IV**
- **b. I and IV**
- **c. I**
- **d. II and III**

Question **2**

Complete

Mark 1.00 out of 1.00

Which of the following method is NOT appropriate to check the association between two continuous variables?

Select one:

- **a. Pearson's correlation**
- **b. Regression**
- **c. T-tes**
- **d. Chi-square test**

Question 3

Complete

Mark 1.00 out of 1.00

Which of the following is/ are TRUE with respect to Pearson's correlation( )?

Select one or more:

- [x] **a. r = 0 means there is no correlation between x and y**
- [ ] **b. r takes a value between 0 to +1**
- [x] **c. Cannot be applied to ordinal variables**
- [x] **d. A measure for the degree of linear dependence between two numerical variables**

Question 4

Complete

Mark 1.00 out of 1.00

Data was collected from a laptop manufacturer for a study, to understand the impact of battery life of a laptop on customer satisfaction (in terms of ratings). The data consisted of ratings (y) provided for each laptop based on their battery life (x) (in years). The following regression model was obtained

$y=1.98+0.5x$

From the given linear regression model, the value 1.98 represents the ___

Select one:

- ( ) **a. error**
- ( ) **b. prediction**
- (•) **c. intercept**
- ( ) **d. slope**

Question 5

Complete

Mark 1.00 out of 1.00

A predictive modeling problem where the class label is predicted for the input data is a type of

Select one:

- ( ) **a. Function approximation problem**
- (•) **b. Classification problem**

Question 6

Complete

Mark 1.00 out of 1.00

Consider the following confusion matrix to answer 26 and 27

		Actual	
		Accept	Reject
Predicted	Accept	15	5
	Reject	1	5

The accuracy of the model is (rounded off to two decimal places): -

Select one:

- **a. 0.71**
- **b. 0.65**
- **c. 0.77**
- **d. None of the mentioned**

Question 7

Complete

Mark 1.00 out of 1.00

The values of the stationary points $x*$,$y*$ and $\lambda*$ for the objective function $f(x,y)=5x-3y$ subject to the given constraints $x2+y2=136$ are.

Select one:

- **a. $x*=\pm10, y*=\pm6$ and $\lambda*=\pm0.25$**
- **b. $x*=y*=\lambda*=-10$**
- **c. None of the mentioned**
- **d. $x*=\pm9, y*=\pm5$ and $\lambda*=\pm0.5$**

Question 8

Complete

Mark 1.00 out of 1.00

The function $f(x,y)=2x^2-2y^2$

Select one:

- **a. has a stationary point at (0,0)**
- **b. has a stationary point at (1,-1)**
- **c. has no stationary point**
- **d. has a stationary point at (1,1)**

Question 9

Complete

Mark 1.00 out of 1.00

For a function $f(x)=3x^4-4x^3-12x^2+45$, which of the following are stationary points and minimisers of f(x)

Select one:

- **a. 0.25,2**
- **b. 0,-2.5**
- **c. 2, -2.5**
- **d. 2, -1**

Question 10

Complete

Mark 1.00 out of 1.00

Which command is used to build a logistic regression model in R?

Select one:

- **a. lm()**
- **b. glm ()**
- **c. None of the mentioned**
- **d. glr ()**

Question 11

Complete

Mark 1.00 out of 1.00

The graph for the straight-line y=mx+b (m - slope, b - intercept) will have a upward slope when the value of

Select one:

- **a. m>0**
- **b. m<0**
- **c. m≥0**
- **d. m=0**

Question 12

Complete

Mark 1.00 out of 1.00

The Eigenvalues of Hessian matrix of is $f(x,y)=2x^2-xy+y^2-3x-y$

Select one:

- **a. 3.828427, -1.828427**
- **b. -1.585786, -4.414214**
- **c. 4.414214, 1.585786**
- **d. -3.828427, 1.828427**

Question 13

Complete

Mark 0.00 out of 1.00

The sensitivity pertaining to the given confusion matrix is (rounded off to two decimal places)

Select one:

- **a. 0.82**
- **b. None of the mentioned**
- **c. 0.71**
- **d. 0.94**

Question 14

Complete

Mark 1.00 out of 1.00

K-Means Clustering model becomes better as

Select one:

- a. we increase the within-cluster SS and increase the between-cluster SS
- b. we decrease the within-cluster SS and increase the between-cluster SS
- c. None of the mentioned
- d. we increase the within-cluster SS and decrease the between-cluster SS

Question 15

Complete

Mark 0.00 out of 1.00

Which of the following is one of the key data science skills?

Select one:

- a. Statistics
- b. Data Visualization
- c. All of the mentioned
- d. Machine Learning

Question 16

Complete

Mark 1.00 out of 1.00

Which of the following statements is/are NOT TRUE?

I. Spearman rank correlation can be used for ordinal variables

II. Pearson's correlation takes a value between 0 to +1

III. A pair of observations (x1,y1) and (x2,y2) that follow the property of x1>x2,y1>y2 or x1<x2,y1<y2 are called concordant pairs

Select one:

- a. I
- b. I and III
- c. III
- d. II and III

Question 17

Complete

Mark 0.00 out of 1.00

Which of the following term(s) is/are associated with K-means clustering?

Select one or more:

- [x] **a. Cluster centers**
- [x] **b. Elbow method**
- [ ] **c. Dendrogram**
- [x] **d. Distance metric**

Question 18

Complete

Mark 1.00 out of 1.00

The Euclidean distance between the two data points X (-4,4) and Y (8,8) is _____ (Rounded off to three decimal places)

Select one:

- ( ) **a. (Type: Range) 9.123,8.235**
- (•) **b. (Type: Range) 12.645,12.650**
- ( ) **c. (Type: Range) 10.655,11.352**
- ( ) **d. (Type: Range) 4.253,6.874**

Question 19

Complete

Mark 1.00 out of 1.00

If $f(x)=5x^4-30x^3+40x^2-60$, then the first order necessary condition for either maxima or minima of $f(x)$ is

Select one:

- ( ) **a. $20x^3-90x^2-80=0$**
- ( ) **b. $20x^3-80x^2-90x=0$**
- ( ) **c. $20x^2-90x^2+80=0$**
- (•) **d. $20x^3-90x^2+80x=0$**

Question **20**

Complete

Mark 1.00 out of 1.00

he condition resulting in non-uniform error variance in regression analysis is termed as:-

Select one:

- **a. Heteroscedasticity**
- **b. Homoscedasticity**

Question **21**

Complete

Mark 1.00 out of 1.00

The Hessian matrix of is $f(x,y)=2x^2-xy+y^2-3x-y$

Select one:

- **a. negative definite**
- **b. positive definite**
- **c. positive semidefinite**
- **d. negative semidefinite**

Question **22**

Complete

Mark 0.00 out of 1.00

The equation for simple linear regression equation is represented as $yi=\beta0+\beta1xi$. The term $\beta1$ denotes:-

Select one:

- **a. Error**
- **b. Estimated intercept**
- **c. Estimated slope**
- **d. Predicted response**

Question 23

Complete

Mark 1.00 out of 1.00

The size of each cluster is (the order of values in each option could be different):-

Select one:

- a. 13 17 20
- b. 11 18 25
- c. 10 17 20
- d. 9 11 18

Question 24

Complete

Mark 1.00 out of 1.00

The Pearson correlation coefficient for the given data is: -

X	5.1	4.9	4.7	4.6	5.0
Y	3.5	3.0	3.2	3.1	3.6

Select one:

- a. 0.68
- b. 0.00
- c. -0.68
- d. 1.00

Question 25

Complete

Mark 1.00 out of 1.00

The most commonly used distance metric to calculate distance between centroid of each cluster and data points in K-means algorithm is

Select one:

- a. Euclidean
- b. None of the mentioned
- c. Manhattan
- d. Chebyshev

Question **26**

Complete

Mark 1.00 out of 1.00

Which method of analysis does not need an independent and dependent variable(s)?

Select one:

- **a. Cluster analysis**
- **b. Discriminant analysis**
- **c. Analysis of variance**
- **d. Regression analysis**

Question **27**

Complete

Mark 1.00 out of 1.00

Point out the correct statement.

Select one:

- **a. None of the mentioned**
- **b. Preprocessed data is original source of data**
- **c. Raw data is original source of data**
- **d. Raw data is the data obtained after processing steps**

Question **28**

Complete

Mark 1.00 out of 1.00

The within cluster sum of squares for each cluster (the order of values in each option could be different):-

Select one:

- **a. 8.561523 10.39874 35.26852**
- **b. 11.95246 19.62285 46.74796**
- **c. 12.35869 18.36987 40.36756**
- **d. 10.62285 15.64962 43.25469**

Question 29

Complete

Mark 1.00 out of 1.00

We intend to find the maxima of f(x,y)=3x+y subject to the constraint x2+y2=10. The Lagrangian function is

Select one:

- a. $L(x,y,\lambda)=\lambda(3x+y-10)$
- b. None of the mentioned
- c. $L(x,y,\lambda)=3x+y$
- d. $L(x,y,\lambda)=3x+y+\lambda(10-x2-y2)$

Question 30

Complete

Mark 1.00 out of 1.00

The Logistic regression tends to overfit when we have large number of independent variables present.

Select one:

- a. False
- b. True

Question 31

Complete

Mark 1.00 out of 1.00

In the simple linear regression equation , where the value 25.9 refers to

Select one:

- a. slope
- b. prediction
- c. intercept
- d. error

Question 32

Complete

Mark 1.00 out of 1.00

The values of the stationary points $x*,y*$ and $\lambda*$ for the objective function $f(x,y)=3x+y$ subject to the given constraints $x2+y2-10=0$ are.

Select one:

- **a. $x*=\pm3, y*=\pm1$ and $\lambda*=\pm0.5$**
- **b. $x*=\pm2, y*=\pm1$ and $\lambda*=\pm5$**
- **c. None of the mentioned**
- **d. $x*=y*=\lambda*=-2$**

Question 33

Complete

Mark 0.00 out of 1.00

Elbow plot can be used to decide the optimal k value in both kNN and K-means clustering problems.

Select one:

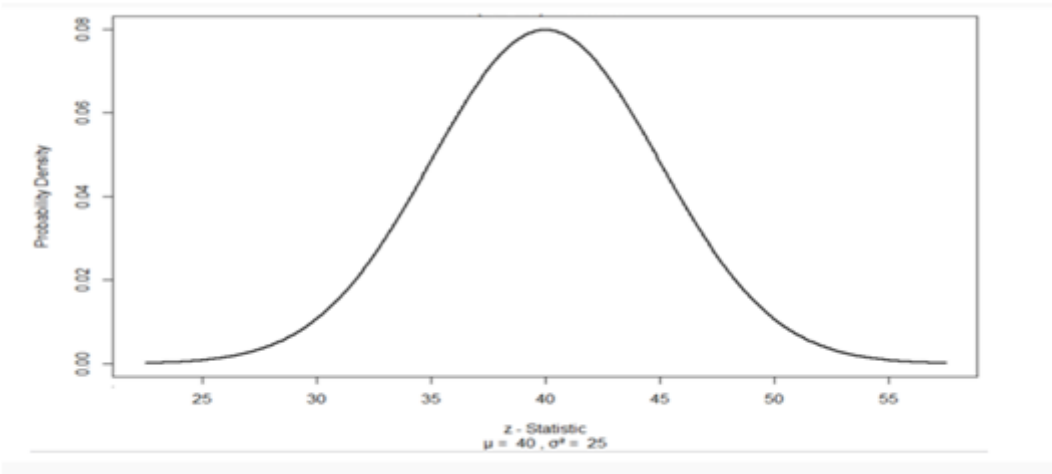
- **a. False**
- **b. True**

Question **34**

Complete

Mark 1.00 out of 1.00

A normal distribution has a mean of 40 and standard deviation of 5.



68% of the distribution can be found between what two numbers?

Select one:

- **a. 0 and 45**
- **b. 30 and 50**
- **c. 35 and 45**
- **d. 0 and 68**

Question **35**

Complete

Mark 1.00 out of 1.00

hat is the Total Sum-of-Squares value of the k-means model? (Choose the appropriate range)

Select one:

- **a. 3000-3200**
- **b. 2600-2800**
- **c. None of the mentioned**
- **d. 2800-3000**

Question 36

Complete

Mark 0.00 out of 1.00

$h(\bar{x})=0$ is

Select one:

- a. Hessian matrix
- b. Inequality constraint
- c. Objective function
- d. Equality constraint

Question 37

Complete

Mark 1.00 out of 1.00

Homoscedasticity in regression analysis is the condition in which the

Select one:

- a. error variance is non-uniform
- b. error variance is equal to zero
- c. error variance remains the same
- d. None of the mentioned

Question 38

Complete

Mark 1.00 out of 1.00

If the objective function $f(\bar{x})$ is quadratic function and the contraints $h(\bar{x}),g(\bar{x})$ are linear, then the type of optimization problem is

Select one:

- a. Non linear programming
- b. Linear programming
- c. Stochastic programming
- d. Quadratic programming

Question **39**

Complete

Mark 1.00 out of 1.00

In the value of SSE is equal to zero, then

Select one:

- **a. correlation is equal to 0**
- **b. correlation is less than 0**
- **c. coefficient of determination(R) is equal to 0**
- **d. coefficient of determination(R) must be equal to 1**

Question **40**

Complete

Mark 1.00 out of 1.00

$f(\bar{x})$ is

Select one:

- **a. Decision variable**
- **b. Hessian matrix**
- **c. Objective function**
- **d. Constraint**

Question **41**

Complete

Mark 1.00 out of 1.00

Which of the following is performed by Data Scientist?

Select one:

- **a. Define the question**
- **b. Challenge results**
- **c. All of the mentioned**
- **d. Create reproducible code**

Question **42**

Complete

Mark 1.00 out of 1.00

In regression analysis, which of the following R2 values indicate a strong linear relationship?

Select one:

- **a. 0.00**
- **b. -0.86**
- **c. 0.90**
- **d. 0.30**

Question **43**

Complete

Mark 1.00 out of 1.00

An ROC curve is plotted between.

Select one:

- **a. (1 – Sensitivity) and Specificity**
- **b. None of the mentioned**
- **c. Sensitivity and Specificity**
- **d. Sensitivity and (1 – Specificity)**

Question **44**

Complete

Mark 1.00 out of 1.00

The coefficient of determination formula is:-

Select one:

- **a. 1 – (SSE / SSR)**
- **b. (SST / SSE) - 1**
- **c. (SSR / SST) – 1**
- **d. 1 – (SSE / SST)**

Question **45**

Complete

Mark 1.00 out of 1.00

The function min f(x,y)=3x+y subject to the given constraints x2+y2<10 is an example of

Select one:

- **a. Multivariate optimization with inequality constraint**
- **b. Unconstrained multivariate optimization**
- **c. None of the mentioned**
- **d. Multivariate optimization with equality constraint**

Question **46**

Complete

Mark 1.00 out of 1.00

In an unconstrained multivariate optimization problem, if the objective function is then the first order derivative is called

Select one:

- **a. gradient**
- **b. hessian**
- **c. polarization**
- **d. all the mentioned**

Question **47**

Complete

Mark 1.00 out of 1.00

For a function $f(x)=5x^4-30x^3+40x^2-60$, which of the following value is a stationary point of f(x)

Select one or more:

- **a. 3.28**
- **b. 0**
- **c. 4**
- **d. 0.2**

Question **48**

Complete

Mark 1.00 out of 1.00

The eigen values for the hessian matrix obtained in Q5 are:

Select one:

- **a. −10.944272, 6.944272**
- **b. 6.45258, 3.556886**
- **c. None of the mentioned**
- **d. 10.472136, 2.527864**

Question **49**

Complete

Mark 1.00 out of 1.00

For a function , the stationary point is

(Hint: Stationary point is a solution to the first order necessary conditions for maxima or minima of

Select one:

- **a. (-1,0)**
- **b. (1,1)**
- **c. (1,0)**
- **d. (0,1)**