

Started on Monday, 29 April 2024, 10:33 PM

State Finished

Completed on Monday, 29 April 2024, 11:03 PM

Time taken 29 mins 59 secs

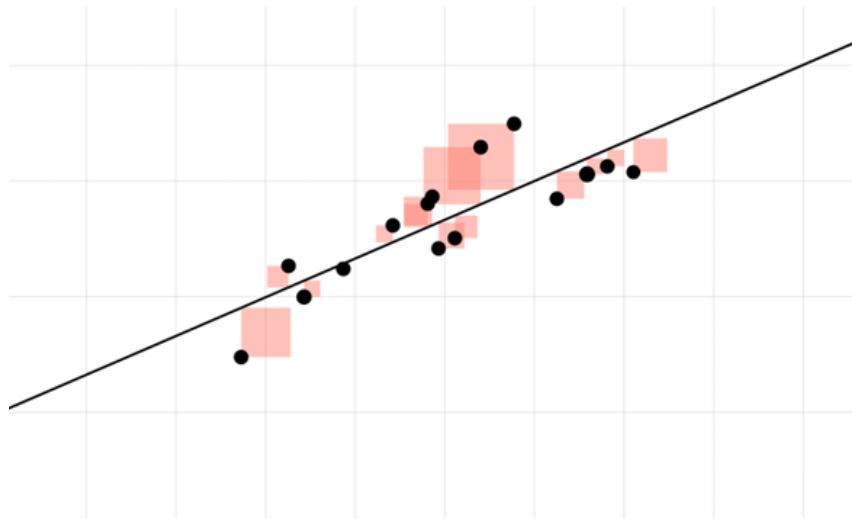
Grade 7.00 out of 10.00 (70%)

**Question 1**

Correct

Mark 1.00 out of 1.00

Given the image below, what do you understand by the size of squares?



- a. None of the options
- b. The greater the size of the square, the better is the model
- c. It is the representation of sum of total squares
- d. Smaller the size of square, better is the line fit ✓

The correct answer is:

Smaller the size of square, better is the line fit

**Question 2**

Correct

Mark 1.00 out of 1.00

Choose the options which represent the wrong assumptions of the logistic regression?

- a. A linear relationship exists between independent variables and the logit of the Target variable.
- b. A linear relationship exists between Independent variables and the Target variable.
- c. A Linear relationship exists between independent variables.
- d. There is minimal or no multicollinearity among the independent variables.
- e. There should be a non-linear relationship between the logit of the outcome and each Independent variable.
- f. The Target variable is in binary.
- g. The independent variables need to be independent of each other.
- h. Usually logistic regression requires a larger dataset to make predictions.

- a. c, e and d
- b. b, c and e ✓
- c. a, f and d
- d. a, c and g

The correct answer is: b, c and e



**Question 3**

Correct

Mark 1.00 out of 1.00

What is true about outliers in logistic Regression ?

- a. Outliers on the correct side of the hyperplane, positively impacts the performance of the model
- b. Outliers on the correct side of the hyperplane, negatively impacts the performance of the model
- c. Outliers on the wrong side of the hyperplane, positively impacts the performance of the model
- d. Outliers on the wrong side of the hyperplane, negatively impacts the performance of the model ✓

The correct answer is: Outliers on the wrong side of the hyperplane, negatively impacts the performance of the model

**Question 4**

Incorrect

Mark 0.00 out of 1.00

Consider the data collected from 500 customers in a restaurant.

It is observed that 40 of the 80 customers tipped the server who was wearing a black jacket, and 300 of the 420 customers tipped the server who was wearing a different color. Compute the \*\*logit (log of odds)\*\* of tipping a server wearing a Red jacket.

- a. 0.723 ✗
- b. 0.174
- c. 0.915
- d. 0.500

The correct answer is: 0.915

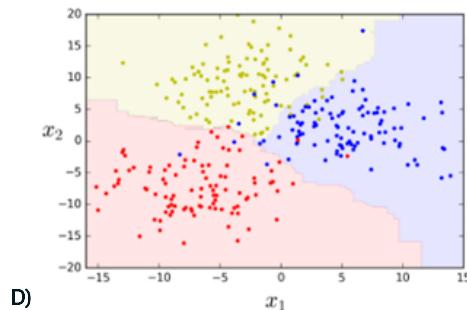
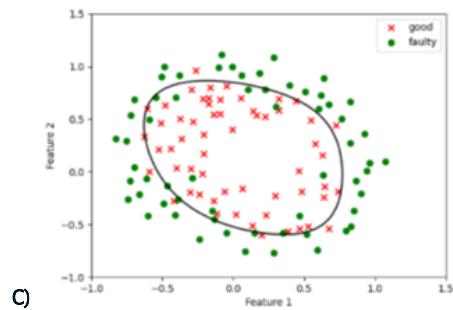
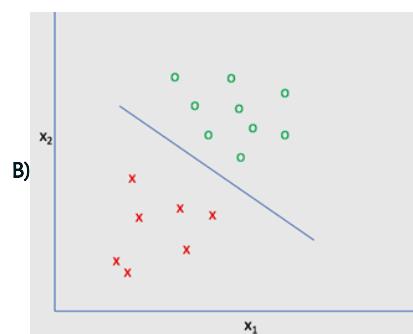
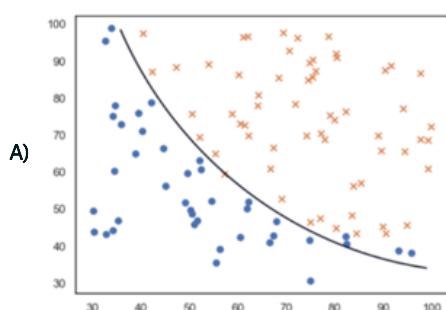


**Question 5**

Correct

Mark 1.00 out of 1.00

Given below are the visualizations of decision boundaries. Which of the following cannot be a decision boundary when a single Logistic regression model is used for classification?



- a. D ✓
- b. B
- c. C
- d. A

The correct answer is: D

**Question 6**

Incorrect

Mark 0.00 out of 1.00

Why do we square the residuals in the Mean squared error (MSE) instead of using modulus?

- a. Squaring the residuals makes sure that negative residuals can't offset the positive residuals. ✗
- b. We want to penalize the predictions that are producing more errors.
- c. Squared error is not differentiable at all points (undefined at 0).
- d. Absolute error is not differentiable at all points (undefined at 0).

The correct answer is: Absolute error is not differentiable at all points (undefined at 0).

**Question 7**

Correct

Mark 1.00 out of 1.00

Based on the socio-economic status of doctors, a researcher wants to build a simple linear regression model to find out if the doctors work at a private or government hospital. Select the most suitable option.

- a. Because there are not enough variables for the analysis, it can't be done.
- b. Socio-economic status cannot be used as an independent variable.
- c. The researcher can make predict using a linear regression model in this scenario.
- d. The target variable is either 0 or 1 and not continuous, hence Linear Regression wont work. ✓

The correct answer is: The target variable is either 0 or 1 and not continuous, hence Linear Regression wont work.

**Question 8**

Correct

Mark 1.00 out of 1.00

How do you check independence of errors in Linear regression?

- a. scatter plot between the target variable and the error term ✓
- b. Scatter plot between various features
- c. Line plot between the target variable and the error term
- d. plot between X and Y

The correct answer is: scatter plot between the target variable and the error term



**Question 9**

Incorrect

Mark 0.00 out of 1.00

Select the correct statement(s) that hold true for regression analysis.

- a. A regression line is created by minimizing the absolute value of the point on the line and the data points. ✗
- b. A regression line is created by minimizing the product of predicted output and expected output.
- c. None of the above.
- d. A regression line is created by minimizing the product of predicted output and expected output.

The correct answer is: A regression line is created by minimizing the product of predicted output and expected output.

**Question 10**

Correct

Mark 1.00 out of 1.00

A farmer wants to know the total amount of crop yield with the help of two features:

1. Amount of Water
2. Amount of Fertilizer

What would be the Null hypothesis for a linear regression model to predict the crop yield with the above-given features?

Note: Usually Null Hypotheses propose that no statistical significance exists in a set of given observations.

- a. There is no linear relationship of crop yield with amount of water and amount of fertilizer. ✓
- b. All the weights of the model are non-zero.
- c. None of the above.
- d. There is linear relationship of crop yield with amount of water and amount of fertilizer.

The correct answer is: There is no linear relationship of crop yield with amount of water and amount of fertilizer.

