X


(https://swayam.gov.in)


(https://swayam.gov.in/nc_details/NPTEL)

Unit 9 - Week 7

Assignment 7

The due date for submitting this assignment has passed. Due on 2020-03-18, 23:59 IST.
As per our records you have not submitted this assignment.

Click here (https://drive.google.com/open?id=1zOrZNhZUk1W7az9KSYlw1uSiMmomZJCS) to download the Data sets

1) The method used to estimate the parameters of logistic regression is ***1 point***

- ordinary least squares
- maximum likelihood estimation
- poisson distribution
- negative binomial distribution

No, the answer is incorrect.
Score: 0
Accepted Answers:
maximum likelihood estimation

2) The value of the sigmoid function when x = +∞ is ***1 point***

- +∞
- -∞
- 0
- 1

No, the answer is incorrect.
Score: 0
Accepted Answers:
1

3) Which of the following measures cannot be used to calculate specificity? ***1 point***

- False Positive
- False Negative
- True Negative
- None of the above

No, the answer is incorrect.
Score: 0
Accepted Answers:
False Negative

4) ROC curve is a plot of ***1 point***

- Precision vs recall
- Accuracy vs misclassification rate
- Specificity vs Sensitivity
- True Positive Rate vs False Positive Rate

No, the answer is incorrect.
Score: 0
Accepted Answers:
True Positive Rate vs False Positive Rate

Based on the information given below answer the questions from **5 to 7**.
Read the given dataset '**auto_mpg.csv'** in Rstudio.

The data is about vehicle performance. The objective is to predict the miles per gallon given the other attributes.

Variables	Description
mpg	Miles per gallon of the engine
cylinders	number of cylinders in the engine
displacement	engine displacement
horsepower	horsepower of the car
weight	weight of the car (lbs)
acceleration	acceleration of the car
model_year	model year of the car in the 1900s

Build the linear regression model '***model_1***' using the above variables

5) The intercept of the full model is (rounded off to two decimal places) ***1 point***

- 1.43
- 0.15
- 7.12
- 4.98

No, the answer is incorrect.
Score: 0
Accepted Answers:
7.12

6) The coefficient of the variable **model_year** is (rounded off to three decimal places) ***1 point***

- 0.459
- -0.041
- 0.005
- 7.123

No, the answer is incorrect.
Score: 0
Accepted Answers:
0.459

7) Which of the following variables are significant at 1% level of significance? ***1 point***

- weight and cylinder
- weight and model_year
- model_year and acceleration
- model_year and horsepower

No, the answer is incorrect.
Score: 0
Accepted Answers:
weight and model_year

Based on the question 7, answer questions 8 and 9.

After identifying the significant variables from question 7, build the new model ***'model_2'*** using the same.

8) The coefficient of determination for ***model_2*** is ***1 point***

- 0.802
- 0.782
- 0.79
- 0.814

No, the answer is incorrect.
Score: 0
Accepted Answers:
0.814

9) The standard error corresponding to the coefficient of **model_year** for **model_2** is (rounded off to three decimal places) ***1 point***

- 0.057
- 0.001
- 0.470
- 4.342

No, the answer is incorrect.
Score: 0
Accepted Answers:
0.057

Build the linear regression model '***model_3***' using all the variables except displacement and acceleration.
Answer questions **10 to 11** based on this model.

10) The adjusted R^2 of **model_3** is **1 point**

- 0.801
- 0.791
- 0.813
- 0.816

No, the answer is incorrect.
Score: 0
Accepted Answers:
0.813

11) The t value corresponding to the coefficient for the variable **horsepower** is **1 point**

- 1.335
- -1.373
- -0.336
- -0.002

No, the answer is incorrect.
Score: 0
Accepted Answers:
-0.336

12) Which of the following performance measures cannot be used for regression? **1 point**

- R-Squared
- RMSE
- Accuracy
- Durbin-Watson

No, the answer is incorrect.
Score: 0
Accepted Answers:
Accuracy

Based on the information given below answer the questions from **14 to 20.**
Given the datasets "**Titanic_Train.csv**" & "**Titanic_Test.csv**" read them as two separate data frames - Train_Data and Test_Data.

The data provides information on the fate of passengers on the fatal maiden voyage of the ocean liner 'Titanic', and the description
of variables are given below. Build a logistic regression model using the **"Train_Data"** after identifying the output class label. Predict the model on the **"Test_Data"** and answer the questions from **14 to 19**.

Variables	Description
Survived	Survived or not (0 = No; 1 = Yes)
Pclass	Passenger Class (1 = 1st; 2 = 2nd; 3 = 3rd)
Sex	Passenger's Sex
Age	Passenger's Age

13) The estimate value of intercept is (rounded off to two decimal places) **1 point**

- 3.95
- 4.69

- 5.83
- 6.69

No, the answer is incorrect.
Score: 0
Accepted Answers:
4.69

14) The standard error for the coefficient of variable **Pclass** is (rounded off to two decimal places) ***1 point***

- 0.86
- 0.36
- 0.16
- 0.57

No, the answer is incorrect.
Score: 0
Accepted Answers:
0.16

15) The accuracy score of the model is ***1 point***

- 0.74
- 0.86
- 0.69
- 0.79

No, the answer is incorrect.
Score: 0
Accepted Answers:
0.79

16) Number of misclassified samples are ***1 point***

- 63
- 24
- 43
- 19

No, the answer is incorrect.
Score: 0
Accepted Answers:
43

17) The proportion of positive cases that were correctly identified is ***1 point***

- 0.768
- 0.782
- 0.724
- 0.791

No, the answer is incorrect.
Score: 0
Accepted Answers:
0.768

18) Balanced accuracy of the model is (rounded off to three decimal places) ***1 point***

- 0.791
- 0.782
- 0.729

0.7241

No, the answer is incorrect.
Score: 0

Accepted Answers:
0.782

19) Misclassification rate of the model is **1 point**

0.79

0.21

1

0.78

No, the answer is incorrect.
Score: 0

Accepted Answers:
0.21