

X

<https://swayam.gov.in>https://swayam.gov.in/nc_details/NPTEL

sai.doc45@gmail.com ✓

NPTEL (<https://swayam.gov.in/explorer?ncCode=NPTEL>) » Business analytics and data mining Modeling using R (course)



Course outline

How does an NPTEL online course work? ()

Week 0 ()

Week 1 ()

Week 2 ()

Week 3 ()

Week 4 ()

Week 5 ()

● Lecture 21
PREDICTION
PERFORMANCE
(unit?
unit=49&lesson=50)

● Lecture 22
MULTIPLE
LINEAR
REGRESSION
(unit?
unit=49&lesson=51)

Week 5 : Assignment 5

The due date for submitting this assignment has passed.

Due on 2023-03-01, 23:59 IST.

Assignment submitted on 2023-03-01, 19:59 IST

1) Which of the following is correct about prediction performance?

1 point

- ☒ The focus is on reducing root-mean-square error of the continuous output variable
- ☐ The focus is on reducing misclassification error of the continuous output variable
- ☐ Prediction performance is same as classification performance
- ☐ None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

The focus is on reducing root-mean-square error of the continuous output variable

2)

1 point

Which of the following is true about the statements below given?

I: 'Goodness of fit' focuses on how well the model fits the data

II: The model obtained in data mining is certainly same as the one obtained in statistical setting

- ☒ Only statement I is correct
- ☐ Only statement II is correct
- ☐ Both the statements are correct
- ☐ None of the above

Yes, the answer is correct.

● Lecture 23
MULTIPLE
LINEAR
REGRESSION
PART-2 (unit?
unit=49&lesson=52)

● Lecture 24
MULTIPLE
LINEAR
REGRESSION
PART-3 (unit?
unit=49&lesson=53)

● Lecture 25
MULTIPLE
LINEAR
REGRESSION
PART-4 (unit?
unit=49&lesson=54)

● Quiz: Week 5
: Assignment
5
(assessment?
name=127)

○ Solution for
Week 5 :
Assignment 5
(unit?
unit=49&lesson=55)

Week 6 ()

Week 7 ()

Week 8 ()

Week 9 ()

Week 10 ()

Week 11 ()

Week 12 ()

Download
Videos ()

Weekly
Feedback ()

Text
Transcripts ()

Score: 1

Accepted Answers:

Only statement I is correct

3) Which partition is used for the performance evaluation of the data mining model? **1 point**

- ☐ Training partition
☐ Validation partition
☒ Test partition
☐ None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Test partition

4) What does the negative overall average error indicate? **1 point**

- ☐ On an average level, the model is over predicting the observations
☒ On an average level, the model is under predicting the observations
☐ On an average level, the model has no impact on predicting the observations
☐ None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

On an average level, the model is over predicting the observations

5) Consider the following statements and choose the correct option: **1 point**

Assertion: RMSE is one of the most preferred metrics of predictive accuracy

Reasoning: It is computed in the same unit as the outcome variable and similar to the standard error of estimate

- ☒ Assertion is correct and reasoning gives the correct explanation
☐ Assertion is correct but reasoning does not give the correct explanation
☐ Assertion is incorrect but reasoning is correct
☐ Both assertion and reasoning are incorrect

Yes, the answer is correct.

Score: 1

Accepted Answers:

Assertion is correct and reasoning gives the correct explanation

6) Which of the following is true about predictive modelling using multiple linear regression? **1 point**

- ☐ Used to estimate best-fit model
☒ Used to estimate values of outcome variable for new records
☐ Prediction performance is of secondary importance
☐ None of the above

Yes, the answer is correct.

Score: 1

Accepted Answers:

Used to estimate values of outcome variable for new records

7) Which of the following is correct about OLS?

1 point

- ☐ OLS computes the sample estimates which minimises the sum of deviations between predicted and actual values
- ☐ OLS computes the sample estimates which maximises the sum of deviations between predicted and actual values
- ☒ OLS computes the sample estimates which minimises the sum of squared deviation between predicted and actual values
- ☐ OLS computes the sample estimates which maximises the sum of squared deviation between predicted and actual values

Yes, the answer is correct.

Score: 1

Accepted Answers:

OLS computes the sample estimates which minimises the sum of squared deviation between predicted and actual values

8) Which of the following assumptions of multiple linear regression can be overlooked in data mining? **1 point**

- ☐ Noise follows a normal distribution
- ☐ Observations are independent
- ☐ Linear relationship holds true
- ☒ Heteroscedasticity

No, the answer is incorrect.

Score: 0

Accepted Answers:

Noise follows a normal distribution

9) Which of the following statements is incorrect with respect to explanatory modelling?

1 point

- ☒ The final model selected may not have the best predictive accuracy
- ☐ Full sample is used to estimate the final model
- ☐ There is no previously assumed structure
- ☐ None of the above

No, the answer is incorrect.

Score: 0

Accepted Answers:

There is no previously assumed structure

10) Which of the following visualisation techniques can be used for measuring predictive performance of a model? **1 point**

- ☐ Decile chart
- ☐ Histogram
- ☐ Lift curve

☒ Both a and c

Yes, the answer is correct.

Score: 1

Accepted Answers:

Both a and c