Final Course Assignment Quiz

Practice Quiz, 20 questions

Congratulations! You passed!

Next Item

1 / 1 point

1.
Download the file delay-UA-cleaned.csv for the remainder of this assignment.

delay-UA-cleaned.csv

Part 1

Build a linear regression model with ARR_DELAY column as the dependent variable. Find the model with one predictor variable and the highest R-squared. Consider all numerical columns except ARR_DELAY15.

Which variable did you choose?

- DAY_OF_MONTH

Correct
Bravo

- CRS_DEP_TIME
- ACTUAL_ELAPSED_TIME
- DISTANCE

1 / 1 point

2.
What is the estimated intercept coefficient of the model ?

-4.5333

2.9019

Final Course Assignment Quiz

Practice Quiz, 20 questions

Correct
Bravo!

2.4321

9.3847

---

1 / 1 point

3.
What is the estimated slope coefficient of the model ?

7.9118

8.7911

3.1879

1.1879

Correct
Bravo!

---

1 / 1 point

4.
Consider a data transformation of the predictor variable you chose before. Which of the following transformation gives the highest R-squared value?

No transformation

Square

Square root

Correct
Bravo!

Log

Final Course Assignment Quiz

Practice Quiz, 20 questions

1 / 1 point

5.

Build a linear regression model with ARR_DELAY column as the dependent variable and all other variables except ARR_DEL15 as explanatory variables. Note that you need to transform the categorical variables ORIGIN and DEST before building the model. You also should not include all dummy variables you created. Exclude the dummy variables ORIGIN_SFO and DEST_SFO in your analysis.

What is the coefficient of DISTANCE?

- 0.03899
- 0.04899
- 0.05899

Correct
Bravo!

- 0.06899

1 / 1 point

6.

What is the the adjusted r-squared for the models you built in the previous question ?

- 0.1257

Correct
Bravo!

- 0.2570
- 0.5701
- 0.7012

1 / 1 point

Final Course Assignment Quiz

Practice Quiz, 20 questions

7.

Split the dataset into training and validation sets using a 60:40 split. Set the random seed for partition to 12345. [This should be the default value unless you changed it at some point.] How many observations are in the training set?

- 1560
- 2340

Correct

Refer to the following video for a refresher: Week 2, video 7.

- 3900
- 4800

1 / 1 point

8.

Again, Build a linear regression model with ARR_DELAY column as the dependent variable and all other variables except ARR_DEL15 as explanatory variables. Recall that you need to transform the categorical variables ORIGIN and DEST before building the model. You also should not include all dummy variables you created. Exclude the dummy variables ORIGIN_SFO and DEST_SFO in your analysis. What is the RMS Error in Validation Data Scoring?

- 61.96
- 72.33

Correct

Bravo!

- 84.56
- 96.44

1 / 1 point

9.

Part 2

Final Course Assignment Quiz

Build logistic regression model and interpret results.

Practice Quiz, 20 questions

Build a logistic regression model using the full dataset with ARR_DEL15 as the dependent variable and all other variables except ARR_DELAY as explanatory variables. Again you need to transform the categorical variables ORIGIN and DEST before building the model. You also should not include all dummy variables you created. Exclude the dummy variables ORIGIN_SFO and DEST_SFO in your analysis.

What is the estimated intercept coefficient of the model?

- [x] 0.4656

Correct
Bravo!

- [ ] 0.6564
- [ ] 2.4656
- [ ] 0.1121

1 / 1 point

10.

What is the estimated slope coefficient for DISTANCE ?

- [ ] 5.7023
- [ ] 9.0022
- [x] 0.002277

Correct
Bravo!

- [ ] 1.2547

1 / 1 point

11.

What is the multiple R-squared ?

Final Course Assignment Quiz

Practice Quiz, 20 questions

- 2.5678
- 0.1420

Correct
Bravo!

- 0.9989
- 0.0021

---

1 / 1 point

12.

Split the dataset into training and validation set using a 60:40 split. Set the random seed for partition to 12345. [This should be the default value unless you changed it at some point.] Rebuild the model with the training data. Report the precision on the validation set.

- 0.2232
- 0.6681
- 0.6792
- 0.7053

Correct
Bravo!

---

1 / 1 point

13.

By default, XLMiner using a cutoff threshold 0.5. Repeat the previous question with a cutoff threshold 0.3. What is the precision on the validation set ?

- 0.2842

0.3821

Final Course Assignment Quiz

Practice Quiz, 20 questions

0.4321

Correct
Bravo!

0.4728

---

0 / 1 point

14.

Part 3

Build a classification tree and interpret results.

Build a classification tree with ARR_DEL15 as the dependent variable and all other variables except ARR_DELAY as explanatory variables. Treat ORIGIN and DEST as categorical variables in your analysis.

Which of the following variables appear to be the most important variables that explains flight delays according to feature importance score?

CRS_ELAPSED_TIME

DISTANCE

This should not be selected
Refer to feature importance score in the bottom of the output sheet

DAY_OF_MONTH

DAY_OF_WEEK

---

1 / 1 point

15.

Split the dataset into training and validation set using a 60:40 split. Set the random seed for partition to 12345. [This should be the default value unless you changed it at some point.] Rebuild the tree with training data. Note: On parameters panel, check “Prune (Using Validation Set).”

What is the precision on the validation data for the single tree ?

0.616438

Final Course Assignment Quiz

Practice Quiz · 20 questions

Correct Response
Bravo!

---

1 / 1 point

16.
What is the precision on the validation data for boosted tree ? Note: In order to build the boosted tree, use Classify->Ensemble->Boosting.

0.6055

Correct Response
Bravo!

---

1 / 1 point

17.
What is the precision on the validation data for bagged tree ? Note: In order to build the bagged tree, use Classify->Ensemble->Bagging.

0.6875

Correct Response
Bravo!

---

1 / 1 point

18.

Build a neural network with ARR_DELAY as the dependent variable and all other variables except ARR_DEL15 as explanatory variables. Treat ORIGIN and DEST as categorical variables in your analysis.

Final Course Assignment Quiz

Practice Quiz, 20 questions

Split the dataset into training and validation set using a 60:40 split. Set the random seed for partition to 12345. [This should be the default value unless you changed it at some point.]

What is the smallest validation error in the column Validation: RMSE for the automatic neural net built using the training set ? Note: To build automatic neural net, use Predict->Neural Network->Automatic Network. Set neuron weight initialization seed to 12345 in step 2.

228.87

Correct Response

Bravo!

---

1 / 1 point

19.

What is the validation error (RMS Error) for the boosted neural net built using the training set? Use step size 0.3 for Boosting:Common; this is the default value if you did not change it. Note: To build boosted neural net, use Predict->Ensemble->Boosting. Set neuron weight initialization seed to 12345 in step 2.

813.236998555901

Correct Response

Bravo!

---

1 / 1 point

20.

Repeat the previous question. Change the step size for Boosting:Common to 0.03. What is the validation error (RMS Error)? This exercise shows that parameter tuning can have significant impact on the quality of the predictive model.

122.8516864

Correct Response

Bravo!

Final Course Assignment Quiz

Practice Quiz, 20 questions