

Week 3 Application Assignment

Quiz, 13 questions

✓ **Congratulations! You passed!**

Next Item



1 / 1
point

1.

Let's reconsider the customer reward program dataset. In this exercise, you will complete a predictive modeling task where the target variable is binary. Using the following data file for this exercise:

crp_cleandata.xlsx

The dataset also contains a column `IndustryType`, which is created based on the column `Industry` in the raw data. Note that `Industry` has many categories. The analyst who prepared the data chose to combine some categories, which resulted in the column `IndustryType`. `IndustryType` has five categories: Department, Discount, Grocery, Restaurants, Specialty. You can create a set of dummy variables based on `IndustryType` in XLMiner by using the Transform functions.

Part I.

Consider logistic regression models with `Reward` column as the target variable. Fit the model with two indicator variables, one indicating whether a retailer is a discount store (i.e., `IndustryType` is Discount), and the other indicating whether a retailer is a grocery store (i.e., `IndustryType` is Grocery). Report the coefficient estimates in the next three questions. [Hint: After you create the dummy variables, use them as Selected Variables (instead of Categorical Variables) in the first step of Logistic Regression.]

What is the estimated intercept coefficient?

0.5108



Correct Response



1 / 1
point

2.

What is the estimated coefficient for IndustryType_Discount (round the answer to 4 decimal places **i.e. x.xxxx**) ?

Quiz, 13 questions

-0.9627



Correct Response

Bravo!



1 / 1
point

3.

What is the estimated coefficient for IndustryType_Grocery (round the answer to four decimal places **i.e. x.xxxx**) ?

-0.7339



Correct Response

Bravo



1 / 1
point

4.

What is the number of true positives? (Specify a whole number.)

40



Correct Response

Bravo!



1 / 1
point

5.

What is the number of true negatives? (Specify a whole number.)

21



Correct Response

Week 3 Application Assignment

Quiz, 13 questions



1 / 1
point

6.
Part II.

Split the dataset into training and validation sets using a 60:40 split (set the seed for partitioning to 12345; this should be the default value if you have not changed it). [Hint: note that there two Partition buttons in XLMiner ribbon. You should use the Partition->Standard Partition in the Data Mining group.] Report the new coefficient estimates in the next three questions. Use the same two predictor variables as in Part I.

What is the estimated intercept coefficient (round the answer to 4 decimal places **i.e. x.xxxx**) ?

0.6567



Correct Response

Bravo!



1 / 1
point

7.
What is the estimated coefficient for IndustryType_Discount (round the answer up to 4 decimal places **i.e. x.xxxx**) ?

-0.8390



Correct Response

Bravo!



1 / 1
point

8.

What is the estimated coefficient for IndustryType_Grocery (round the answer to 4 decimal places i.e. x.xxx).

Quiz, 13 questions

-1.1675



Correct Response



1 / 1
point

9.

How many observations are in the training set?

60



Correct Response

Bravo!



1 / 1
point

10.

What is the number of true positives on the validation data? (Specify a whole number.)

13



Correct Response

Bravo!



1 / 1
point

11.

What is the number of true negatives on the validation data? (Specify a whole number.)

10

Week 3 Application Assignment

Quiz, 10 questions

Correct Response
Bravo!



1 / 1
point

12.

(Part 3) By default, XLMiner uses the cutoff threshold 0.5. Repeat Part II with a cutoff threshold 0.3. What are the numbers of true positives and true negatives on the validation data?

Report the number of true positives:

20



Correct Response

Bravo!



1 / 1
point

13.

Report the number of true negatives:

0



Correct Response

Bravo!



Week 3 Application Assignment

Quiz, 13 questions