Predicting sentiment from product reviews

13 questions

1.

Are you using GraphLab Create? Please make sure that

1. You are using version 1.8.3 of GraphLab Create. Verify the version of GraphLab Create by running

```
graphlab.version
```

inside the notebook. If your GraphLab version is incorrect, see this post (https://www.coursera.org/learn/ml-classification/supplement/LgZ3I/installing-correct-version-of-graphlab-create) to install version 1.8.3.

2. You are using the IPython notebook named module-2-linear-classifier-assignment-blank.ipynb obtained from the associated reading.

This question is ungraded. Check one of the two options to confirm.

- ○ I confirm that I am using the right version of GraphLab Create and the right IPython notebook.
- ○ I am using another tool, such as scikit-learn.

2.

How many weights are greater than or equal to 0?

68419

3.

Of the three data points in sample_test_data, which one has the lowest probability of being classified as a positive review?

- ○ First
- ○ Second
- ○ Third

4.

Which of the following products are represented in the 20 most positive reviews?

- ☐ Snuza Portable Baby Movement Monitor
- ☐ MamaDoo Kids Foldable Play Yard Mattress Topper, Blue
- ☐ Britax Decathlon Convertible Car Seat, Tiffany
- ☐ Safety 1st Exchangeable Tip 3 in 1 Thermometer

5.

Which of the following products are represented in the 20 most negative reviews?

- ☐ The First Years True Choice P400 Premium Digital Monitor, 2 Parent Unit
- ☐ JP Lizzy Chocolate Ice Classic Tote Set
- ☐ Peg-Perego Tatamia High Chair, White Latte
- ☐ Safety 1st High-Def Digital Monitor

6.

What is the accuracy of the sentiment_model on the test_data? Round your answer to 2 decimal places (e.g. 0.76).

0.91

7.

Does a higher accuracy value on the training_data always imply that the classifier is better?

- (●) Yes, higher accuracy on training data always implies that the classifier is better.
- ( ) No, higher accuracy on training data does not necessarily imply that the classifier is better.

8.

Consider the coefficients of simple_model. There should be 21 of them, an intercept term + one for each word in significant_words.

How many of the 20 coefficients (corresponding to the 20 significant_words and excluding the intercept term) are positive for the simple_model?

10

9.

Are the positive words in the simple_model also positive words in the sentiment_model?

- ( ) Yes
- (●) No

10.

Which model (sentiment_model or simple_model) has higher accuracy on the TRAINING set?

- ( ) Sentiment_model
- (●) Simple_model

11.

Which model (sentiment_model or simple_model) has higher accuracy on the TEST set?

- ○ Sentiment_model
- ● Simple_model

12.

Enter the accuracy of the majority class classifier model on the test_data. Round your answer to two decimal places (e.g. 0.76).

0.84

13.

Is the sentiment_model definitely better than the majority class classifier (the baseline)?

- ○ Yes
- ● No

Submit Quiz