Exploring precision and recall

✔ **14/14** points earned (100%)

Quiz passed!

Continue Course (/learn/ml-classification/supplement/rC1np/slides-presented-in-this-module)

Back to Week 6 (/learn/ml-classification/home/week/6)

✔ 1 / 1
   points

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. **This assignment is not guaranteed to work with other versions of GraphLab Create.**

2. You are using the IPython notebook named module-9-precision-recall-assignment-blank.ipynb obtained from the associated reading.

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

---

1 / 1
points

2.
Consider the logistic regression model trained on **amazon_baby.gl** using GraphLab Create.

Using accuracy as the evaluation metric, was our **logistic regression model** better than the **majority class classifier**?

---

1 / 1
points

3. How many predicted values in the **test set** are **false positives**?

✓ **1 / 1 points**

4.

Consider the scenario where each false positive costs $100 and each false negative $1.

Given the stipulation, what is the cost associated with the logistic regression classifier's performance on the **test set**?

✓ **1 / 1 points**

5.

Out of all reviews in the **test set** that are predicted to be positive, what fraction of them are **false positives**? (Round to the second decimal place e.g. 0.25)

✓ **1 / 1 points**

6.

Based on what we learned in lecture, if we wanted to reduce this fraction of false positives to be below 3.5%, we would:

✓ **1 / 1 points**

7.

What fraction of the positive reviews in the **test_set** were correctly predicted as positive by the classifier? Round your answer to 2 decimal places.

1 / 1 points

8. What is the recall value for a classifier that predicts **+1** for all data points in the **test_data**?

1 / 1 points

9.

What happens to the number of positive predicted reviews as the threshold increased from 0.5 to 0.9?

1 / 1 points

10.

Consider the metrics obtained from setting the threshold to 0.5 and to 0.9.

Does the **precision** increase with a higher threshold?

1 / 1

✓ points

11.

Among all the threshold values tried, what is the **smallest** threshold value that achieves a precision of 96.5% or better? Round your answer to 3 decimal places.

✓ 1 / 1 points

12.

Using threshold = 0.98, how many **false negatives** do we get on the **test_data**? (**Hint**: You may use the graphlab.evaluation.confusion_matrix function implemented in GraphLab Create.)

✓ 1 / 1 points

13.

Questions 13 and 14 are concerned with the reviews that contain the word **baby**.

Among all the threshold values tried, what is the **smallest** threshold value that achieves a precision of 96.5% or better for the reviews of data in **baby_reviews**? Round your answer to 3 decimal places.

✓ 1 / 1 points

14.

Questions 13 and 14 are concerned with the reviews that contain the word **baby**.

Is this threshold value smaller or larger than the threshold used for the entire dataset to achieve the same specified precision of 96.5%?