Congratulations! You passed!

Next Item





You are working on a spam classification system using regularized logistic regression. "Spam" is a positive class (y = 1) and "not spam" is the negative class (y = 0). You have trained your classifier and there are m = 1000 examples in the cross-validation set. The chart of predicted class vs. actual class is:

	Actual Class: 1	Actual Class: 0
Predicted Class: 1	85	890
Predicted Class: 0	15	10

For reference:

- . Accuracy = (true positives + true negatives) / (total examples)
- Precision = (true positives) / (true positives + false positives)
- Recall = (true positives) / (true positives + false negatives)
- F_1 score = (2 * precision * recall) / (precision + recall)

What is the classifier's F_1 score (as a value from 0 to 1)?

Enter your answer in the box below. If necessary, provide at least two values after the

0.1578

Precision is 0.087 and recall is 0.85, so F_1 score is (2 * precision * recall) / (precision + recall) = 0.158.



Suppose a massive dataset is available for training a learning algorithm. Training on a lot of data is likely to give good performance when two of the following conditions hold true.

Which are the two?



large number of parameters (that is able to

learn/represent fairly complex functions).

You should use a "low bias" algorithm with many parameters, as it will be able to make use of the large dataset provided. If the model has too few parameters, it will underfit the large training set.

We train a learning algorithm with a

small number of parameters (that is thus unlikely to

overfit).

When we are willing to include high

order polynomial features of x (such as $x_1^2, x_2^2,$

 x_1x_2 , etc.).

Un-selected is correct

information to predict y accurately. (For example, one

way to verify this is if a human expert on the domain

can confidently predict y when given only x).

amount of data can solve a learning problem in which the features do not contain enough information to make an accurate prediction.





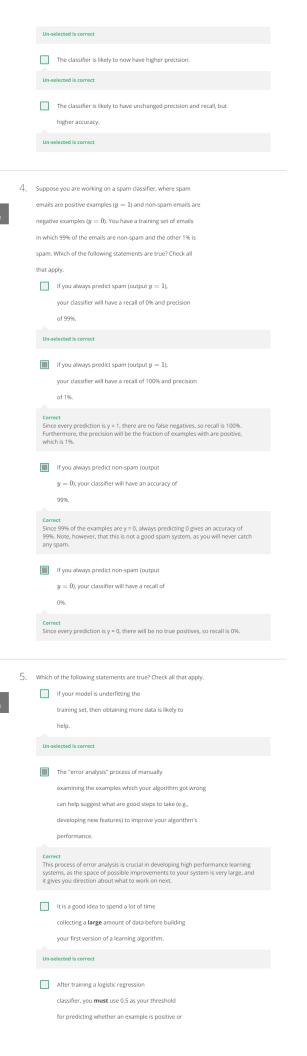
Currently, you predict 1 if $h_{\theta}(x) \geq$ threshold, and predict 0 if $h_{\theta}(x)lt$ threshold, where currently the threshold is set to 0.5. Suppose you decrease the threshold to 0.1. Which of the following are true? Check all

The classifier is likely to now have higher recall.

Lowering the threshold means more y = 1 predictions. This will increase the number of true positives and decrease the number of false negatives, so recall will increase.

The classifier is likely to have unchanged precision and recall, but

lower accuracy.



negative.

Un-selected is correct

Using a **very large** training set

makes it unlikely for model to overfit the training

data.

Correct
A sufficiently large training set will not be overfit, as the model cannot overfit some of the examples without doing poorly on the others.

6 P P