## Machine Learning System Design

Quiz, 5 questions

1 point

1.

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 decimal point.

0.16

1 point Suppose a massive dataset is available for training a learning algorithm. Machine Learning Systematics when two of the following conditions hold true.

Which are the two?

When we are willing to include high			
order polynomial features of $x$ (such as $x_1^2$ , $x_2^2$ ,			
$x_1x_2$ , etc.).			

The features x contain sufficient 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).

We train a learning algorithm with a large number of parameters (that is able to learn/represent fairly complex functions).

We train a learning algorithm with a small number of parameters (that is thus unlikely to overfit).

1 point

3.

Suppose you have trained a logistic regression classifier which is outputing  $h_{\theta}(x)$ .

## Machine Learning System Design

Quiz, 5 questions

Currently, you predict 1 if  $h_{\theta}(x) \geq \text{threshold}$ , and predict 0 if  $h_{\theta}(x) < \text{threshold}$ , where currently the threshold is set to 0.5.

Suppose you **decrease** the threshold to 0.3. Which of the following are true? Check all that apply.

The classifier is likely to now have higher recall.
The classifier is likely to have unchanged precision and recall, but
higher accuracy.
The classifier is likely to have unchanged precision and recall, but
lower accuracy.
The classifier is likely to now have higher precision.

1 point

4.

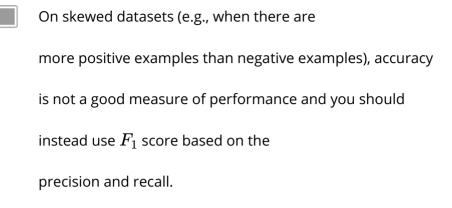
Suppose you are working on a spam classifier, where spam emails are positive examples (y=1) and non-spam emails are negative examples (y=0). You have a training set of emails in which 99% of the emails are non-spam and the other 1% is spam. Which of the following statements are true? Check all that apply.

If you always predict non-spam (output y=0), your classifier will have an accuracy of 99%.

		If you always predict non-spam (output			
Machine Lea	arnin	g Systemclesign have 99% accuracy on the			
Quiz, 5 questions		training set, but it will do much worse on the cross			
		validation set because it has overfit the training			
		data.			
		A good classifier should have both a			
		high precision and high recall on the cross validation			
		set.			
		If you always predict non-spam (output			
		y=0), your classifier will have 99% accuracy on the			
		training set, and it will likely perform similarly on			
		the cross validation set.			
	1 point				
	5.				
	Which	of the following statements are true? Check all that apply.			
		If your model is underfitting the			
		training set, then obtaining more data is likely to			
		help.			
		It is a good idea to spend a lot of time			
		collecting a <b>large</b> amount of data before building			
		your first version of a learning algorithm.			
		After training a logistic regression			

## Machine Learning System Design

$\bigcap$	_	$\alpha$	ınct	ionc
Quiz,	2	uι	ıesı	10115



Using a **very large** training set

makes it unlikely for model to overfit the training

data.



I understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.

Learn more about Coursera's Honor Code

Dhruva Bansal

Submit Quiz





