1 point

1.

You are working on a spam classification system using regularized logistic $Machine\ Learnsippg$ "Syst" employing lass (y = 1) and "not spam" is the negative class Quiz, 5 questions (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 accuracy (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.



2.

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?

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

Machine Lemning System Designwed.

Quiz, 5 questions

A human expert on the application domain	
can confidently predict \boldsymbol{y} when given only the features \boldsymbol{x}	
(or more generally, if we have some way to be confident	
that \boldsymbol{x} contains sufficient information to predict \boldsymbol{y}	
accurately).	
Our learning algorithm is able to	
represent fairly complex functions (for example, if we	
train a neural network or other model with a large	
number of parameters).	

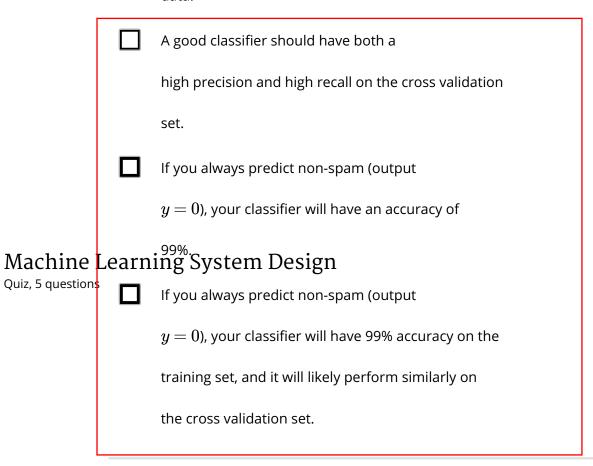
1 point

3.

	Suppose you have trained a logistic regression classifier which is outputing $h_{ heta}(x)$.
	Currently, you predict 1 if $h_{ heta}(x) \geq ext{threshold}$, and predict 0 if $h_{ heta}(x) < ext{threshold}$, where currently the threshold is set to 0.5.
	Suppose you increase the threshold to 0.7. Which of the following are true? Check all that apply.
	The classifier is likely to now have lower precision.
	The classifier is likely to now have lower recall.
Machine L Quiz, 5 questions	emning System Design The classifier is likely to have unchanged precision and recall, but
·	lower accuracy.
	The classifier is likely to have unchanged precision and recall, but
	higher accuracy.
	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 99% accuracy on the training set, but it will do much worse on the cross validation set because it has overfit the training

data.



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.

Using a **very large** training set

makes it unlikely for model to overfit the training

data.

It is a good idea to spend a lot of time

collecting a **large** amount of data before building

Machine Le Quiz, 5 questions	your first version of a learning algorithm.	
	After training a logistic regression	
	chassifier, you must use 0.5 as your threshold	
	for predicting whether an example is positive or	
	negative.	
	The "error analysis" process of manually	
	examining the examples which your algorithm got wrong earning System Design	
	can help suggest what are good steps to take (e.g.,	
	developing new features) to improve your algorithm's	
	performance.	
-		
	I, Jun-Chieh Wang , 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	
	Submit Quiz	



