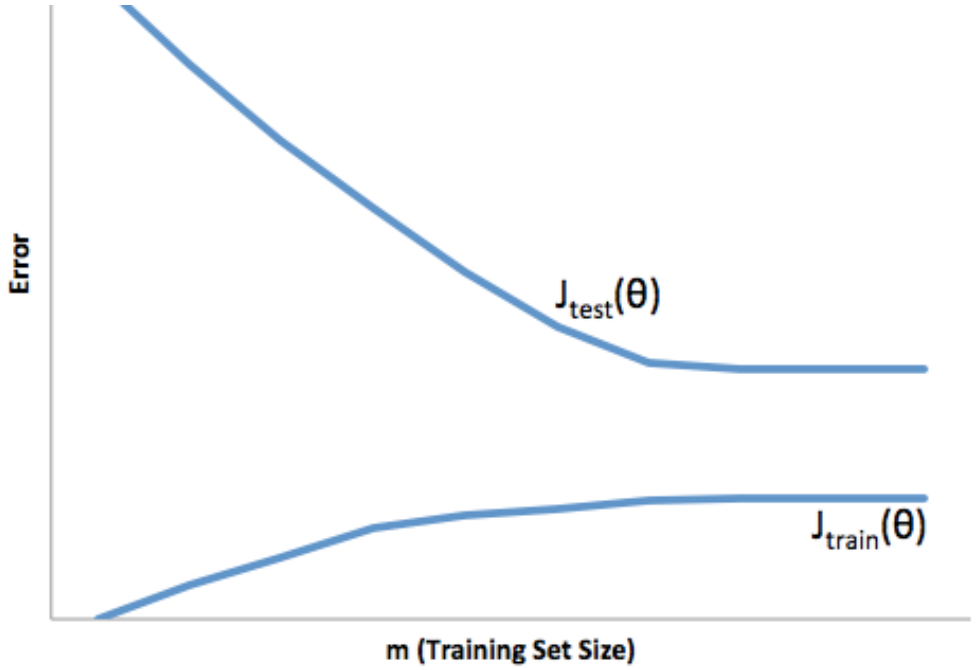
Feedback — X. Advice for Applying Machine Learning

You submitted this quiz on **Fri 31 May 2013 1:57 PM PDT (UTC -0700)**. You got a score of **5.00** out of **5.00**.

Question 1

Does the following learning curve correspond to a learning algorithm that is suffering from high bias, high variance, or neither?



Your Answer		Score	Explanation
High bias			
High	✔	1.00	This learning curve shows a large gap between training and test set

variance		errors, so the algorithm is suffering from high variance.
○ Neither		
Total	1.00 / 1.00	

Question 2

Suppose you have implemented regularized logistic regression to classify what object is in an image (i.e., to do object recognition). However, when you test your hypothesis on a new set of images, you find that it makes unacceptably large errors with its predictions on the new images. However, your hypothesis performs **well** (has low error) on the training set. Which of the following are promising steps to take? Check all that apply.

Your Answer		Score	Explanation
☐ Try adding polynomial features.	✔	0.25	The gap in errors between training and test suggests a high variance problem in which the algorithm has overfit the training set. Using more complex features will only increase the overfitting of the training set.
☑ Try increasing the regularization parameter λ.	✔	0.25	The gap in errors between training and test suggests a high variance problem in which the algorithm has overfit the training set. Increasing the regularization parameter will reduce overfitting and help with the variance problem.
☐ Try evaluating the hypothesis on a cross validation set rather than the test set.	✔	0.25	A cross validation set is useful for choosing the optimal non-model parameters like the regularization parameter λ, but the train / test split is sufficient for debugging problems with the algorithm itself.
☑ Get more training examples.	✔	0.25	The gap in errors between training and test suggests a high variance problem in which the algorithm has overfit

			the training set. Adding more training data will increase the complexity of the training set and help with the variance problem.
Total		1.00 / 1.00	

Question 3

Suppose you have implemented regularized logistic regression to predict what items customers will purchase on a web shopping site. However, when you test your hypothesis on a new set of customers, you find that it makes unacceptably large errors in its predictions. Furthermore, the hypothesis performs **poorly** on the training set. Which of the following might be promising steps to take? Check all that apply.

Your Answer		Score	Explanation
☐ Try using a smaller set of features.	✔	0.25	The poor performance on both the training and test sets suggests a high bias problem. Using fewer features will decrease the complexity of the hypothesis and will make the bias problem worse.
☑ Try decreasing the regularization parameter λ.	✔	0.25	The poor performance on both the training and test sets suggests a high bias problem. Decreasing the regularization parameter will allow the hypothesis to fit the data more closely, improving both training and test set performance.
☐ Try evaluating the hypothesis on a cross validation set rather than the test set.	✔	0.25	A cross validation set is useful for choosing the optimal non-model parameters like the regularization parameter λ, but the train / test split is sufficient for debugging problems with the algorithm itself.
☑ Try to obtain and use additional	✔	0.25	The poor performance on both the training and test sets suggests a high bias problem. Using additional features

features.		will increase the complexity of the hypothesis, thereby improving the fit to both the train and test data.
Total	1.00 / 1.00	

Question 4

Which of the following statements are true? Check all that apply.

Your Answer		Score	Explanation
☐ Suppose you are training a regularized linear regression model.The recommended way to choose what value of regularization parameter λ to use is to choose the value of λ which gives the lowest **training set** error.	✔	0.25	You should not use training error to choose the regularization parameter, as you can always improve training error by using less regularization (a smaller value of λ). But too small of a value will not generalize well on the test set.
☑ Suppose you are using linear regression to predict housing prices, and your dataset comes sorted in order of increasing sizes of houses. It is then important to randomly shuffle the dataset before splitting it into training, validation and test sets, so that we don't have all the smallest houses going into the training set, and all the largest houses going into the test set.	✔	0.25	We want each of the training, cross validation, and test sets to have the same data distribution. Shuffling pre-sorted data ensures this is the case.
☑ A typical split of a dataset into training, validation and	✔	0.25	This is a good split of

test sets might be 60% training set, 20% validation set, and 20% test set.			the data, as it dedicates the bulk of the data to finding model parameters in training while leaving enough data for cross validation and estimating generalization error.
☐ Suppose you are training a logistic regression classifier using polynomial features and want to select what degree polynomial (denoted d in the lecture videos) to use. After training the classifier on the entire training set, you decide to use a subset of the training examples as a validation set. This will work just as well as having a validation set that is separate (disjoint) from the training set.	✔	0.25	This will not work as well as using a separate cross validation set, since the model parameters have already been fit to training data, so using training data for validation will not give an accurate estimate of test set error.
Total		1.00 / 1.00	

Question 5

Which of the following statements are true? Check all that apply.

Your Answer		Score	Explanation
☑ A model with more parameters is more prone to overfitting and typically has higher variance.	✔	0.25	More model parameters increases the model's complexity, so it can more tightly fit data in training, increasing the chances of overfitting.

☑ If a learning algorithm is suffering from high variance, adding more training examples is likely to improve the test error.	✔	0.25	With high variance, the model is overfitting the training data. Adding more training data will increase the complexity of the the train set, thereby reducing the chances of overfitting.
☐ If a neural network has much lower training error than test error, then adding more layers will help bring the test error down because we can fit the test set better.	✔	0.25	With lower training than test error, the model has high variance. Adding more layers will increase model complexity, making the variance problem worse.
☑ When debugging learning algorithms, it is useful to plot a learning curve to understand if there is a high bias or high variance problem.	✔	0.25	The shape of a learning curve is a good indicator of bias or variance problems with your learning algorithm.
Total		1.00 / 1.00	