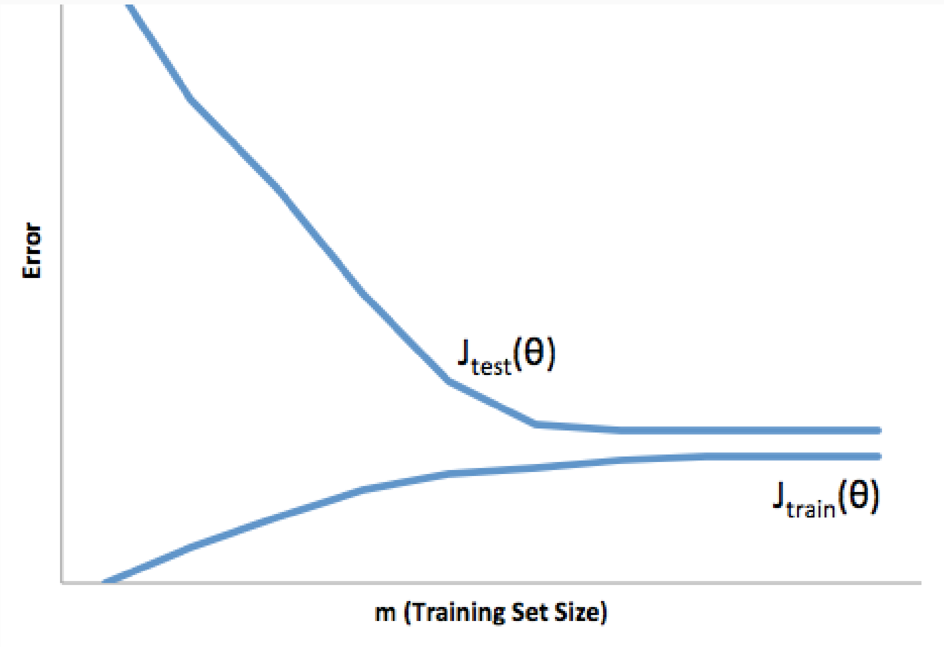
Feedback — X. Advice for Applying Machine Learning

You submitted this quiz on **Sun 27 Apr 2014 5:54 AM IST**. You got a score of **4.25** out of **5.00**. You can attempt again in 10 minutes. Help

Question 1

You train a learning algorithm, and find that it has unacceptably high error on the test set. You plot the learning curve, and obtain the figure below. Is the algorithm suffering from high bias, high variance, or neither?



Your Answer		Score	Explanation
High bias	✔	1.00	This learning curve shows high error on both the training and test sets, so the algorithm is suffering from high bias.
Neither			
High variance			
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 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.
☑ Try adding polynomial features.	✘	0.00	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 using a smaller set of 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. Reducing the feature set will ameliorate the overfitting and help with the variance problem.
Total		0.75 / 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 adding polynomial features.	✔	0.25	The poor performance on both the training and test sets suggests a high bias problem. Adding more complex features will increase the complexity of the hypothesis, thereby improving the fit to both the train and test data.
☑ Try to obtain and use additional features.	✔	0.25	The poor performance on both the training and test sets suggests a high bias problem. Using additional features will increase the complexity of the hypothesis, thereby improving the fit to both the train and test data.
☐ 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 evaluating the hypothesis on a cross validation set rather than the test set.	✖	0.00	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.
Total		0.75 / 1.00	

Question 4

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

Your Answer		Score	Explanation
☑ It is okay to use data from the test set to choose the regularization parameter λ, but not the model parameters (θ).	✖	0.00	You should not use test set data in choosing the regularization parameter, as it means the test error will not be a good estimate of generalization error.
☑ Suppose you are using linear regression to predict housing prices, and your dataset comes sorted in order of increasing sizes of	✔	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.

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.			
☐ 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.
☑ The performance of a learning algorithm on the training set will typically be better than its performance on the test set.	✔	0.25	The learning algorithm finds parameters to minimize training set error, so the performance should be better on the training set than the test set.
Total		0.75 / 1.00	

Question 5

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

Your Answer		Score	Explanation
☐ 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.
☑ If a learning algorithm is suffering from high bias, only adding more training examples may **not** improve the test error significantly.	✔	0.25	With high bias, the model is not fitting the training data currently present, so adding more data is unlikely to help.
☑ 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.
Total		1.00 / 1.00	