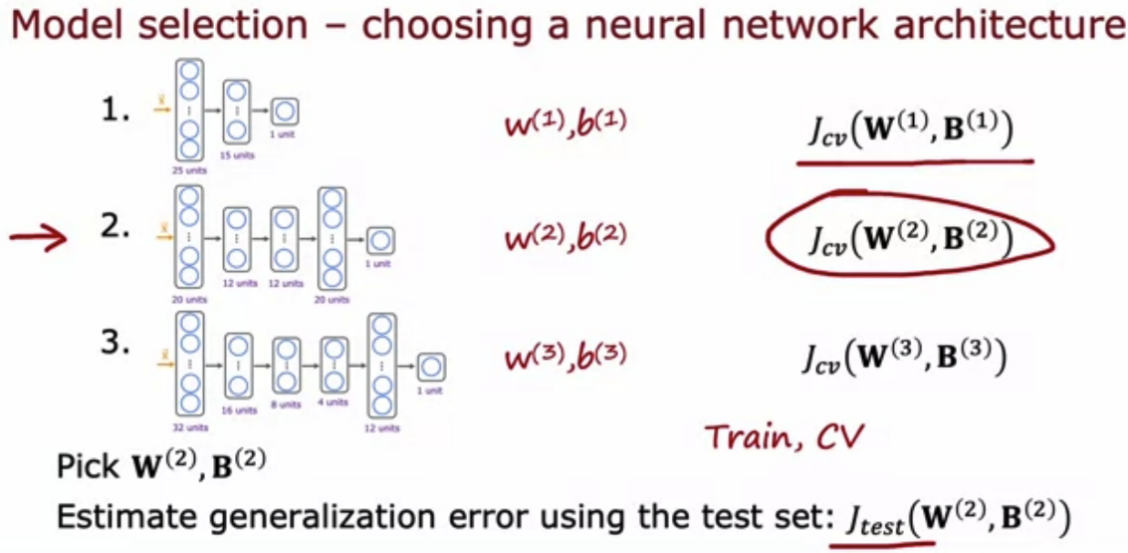
# 第一单元

Advice for applying machine learning

应用机器学习的建议

1. In the context of machine learning, what is a diagnostic? A
2. A test that you run to gain insight into what is/isn’t working with a learning algorithm.
3. A process by which we quickly try as many different ways to improve an algorithm as possible, so as to see what works.
4. This refers to the process of measuring how well a learning algorithm does on a test set (data that the algorithm was not trained on).
5. An application of machine learning to medical applications, with the goal of diagnosing patients’ conditions.
6. True/False? It is always true that the better an algorithm does on the training set, the better it will do on generalizing to new data. B
7. True
8. False



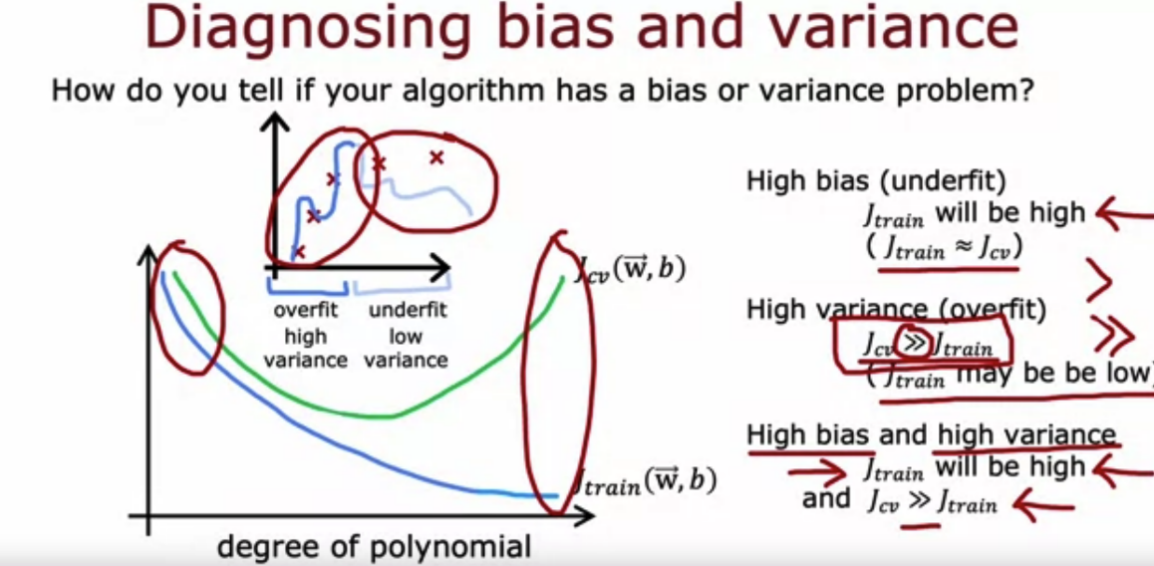
For a classification task; suppose you train three different models using three different neural network architectures. Which data do you use to evaluate the three models in order to choose the best one? B

1. The test set
2. The cross validation set
3. The training set
4. All the data -- training, cross validation and test sets put together.

# 第二单元

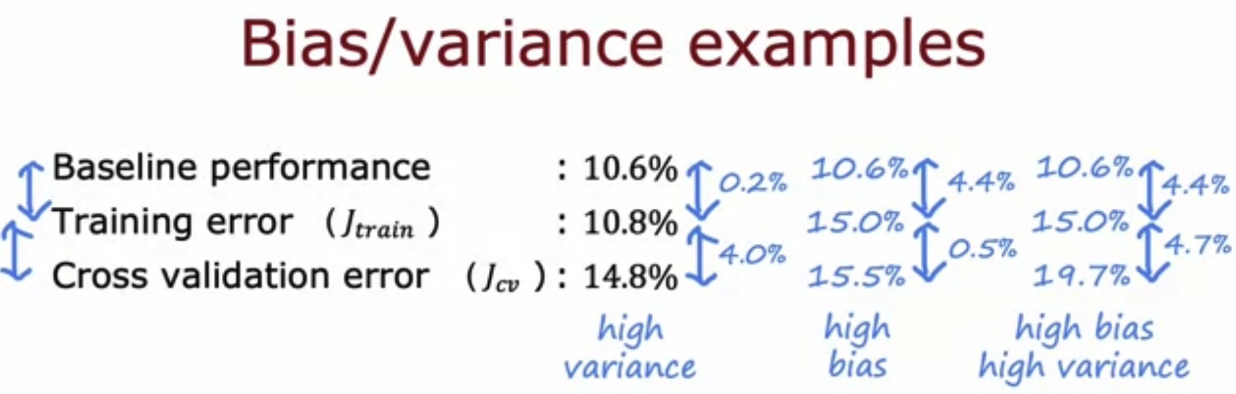
Bias and variance

偏差和方差



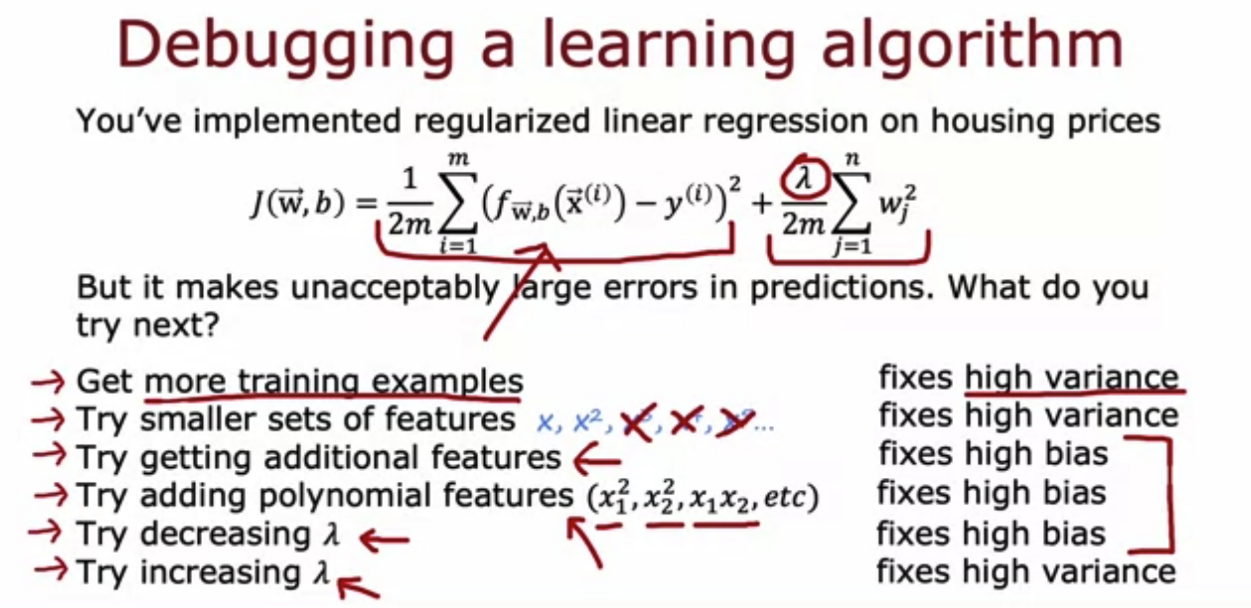
If the model's cross validation error is much higher than the training error , this is an indication that the model has… C

1. Low variance
2. high bias
3. high variance
4. Low bias



Which of these is the best way to determine whether your model has high bias (has underfit the training data)? C

1. See if the cross validation error is high compared to the baseline level of performance
2. See if the training error is high (above 15% or so)
3. Compare the training error to the cross validation error.
4. Compare the training error to the baseline level of performance



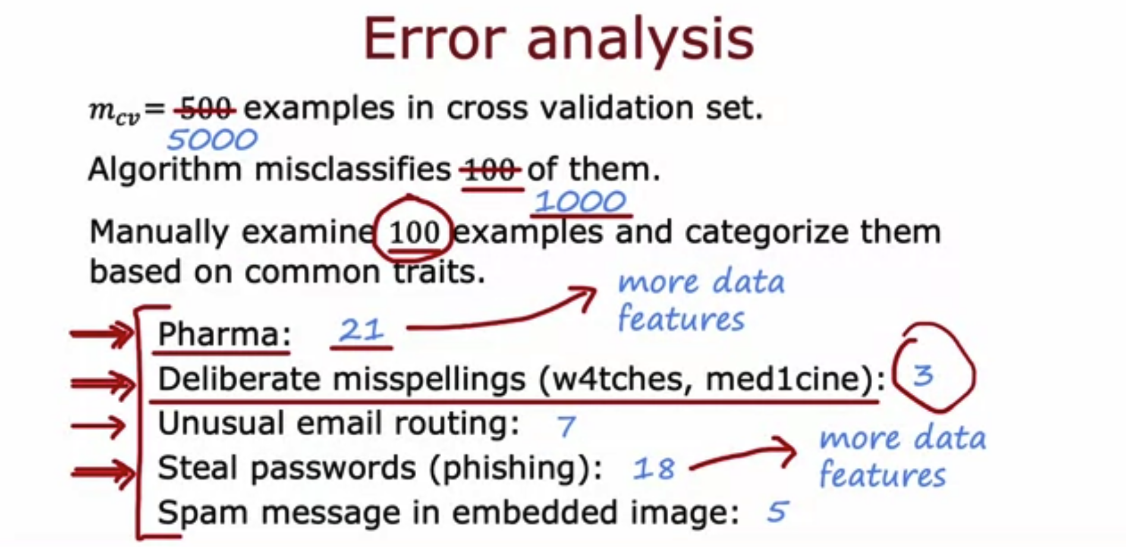
You find that your algorithm has high bias. Which of these seem like good options for improving the algorithm’s performance? Hint: two of these are correct. A B

1. Collect additional features or add polynomial features
2. Decrease the regularization parameter λ (lambda)
3. Collect more training examples
4. Remove examples from the training set
5. You find that your algorithm has a training error of 2%, and a cross validation error of 20% (much higher than the training error). Based on the conclusion you would draw about whether the algorithm has a high bias or high variance problem, which of these seem like good options for improving the algorithm’s performance? Hint: two of these are correct. A C
6. Collect more training data
7. Reduce the training set size
8. Increase the regularization parameter λ
9. Decrease the regularization parameter λ

# 第三单元

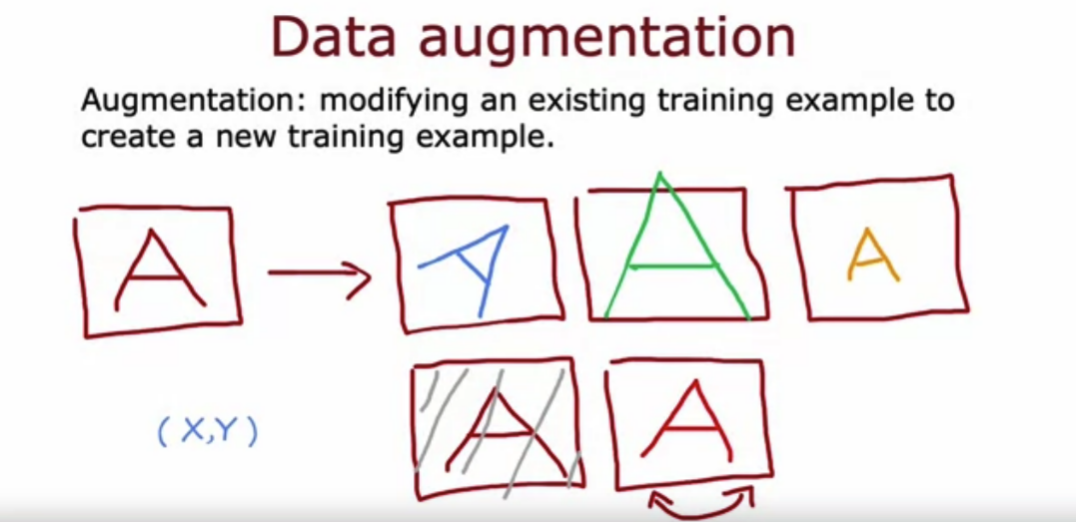
Machine learning development process

机器学习开发流程



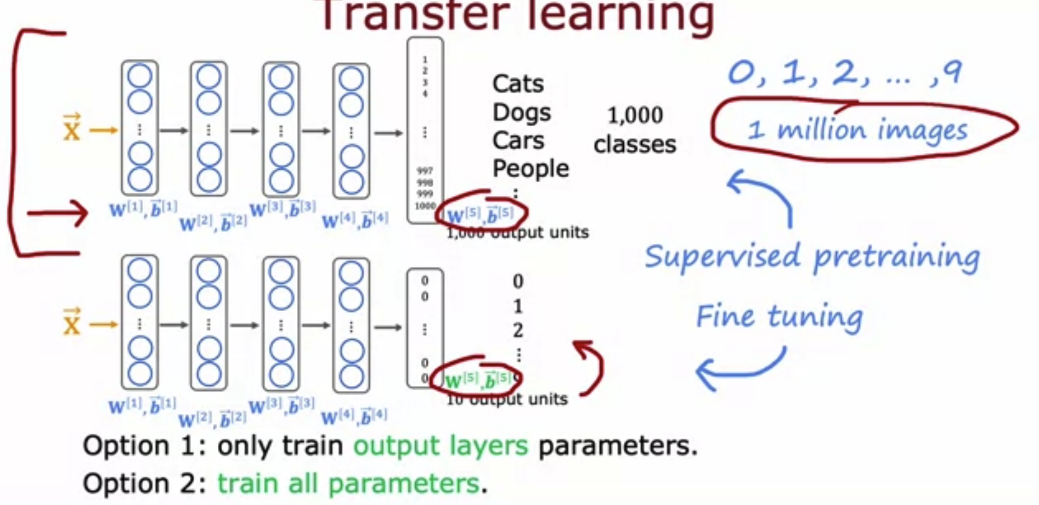
Which of these is a way to do error analysis? D

1. Collecting additional training data in order to help the algorithm do better.
2. Calculating the test error
3. Calculating the training error
4. Manually examine a sample of the training examples that the model misclassified in order to identify common traits and trends.



We sometimes take an existing training example and modify it (for example, by rotating an image slightly) to create a new example with the same label. What is this process called? C

1. Error analysis
2. Machine learning diagnostic
3. Data augmentation
4. Bias/variance analysis



What are two possible ways to perform transfer learning? Hint: two of the four choices are correct. A C

1. You can choose to train just the output layers' parameters and leave the other parameters of the model fixed.
2. Download a pre-trained model and use it for prediction without modifying or re-training it.
3. Given a dataset, pre-train and then further fine tune a neural network on the same dataset.
4. You can choose to train all parameters of the model, including the output layers, as well as the earlier layers.