

Evaluating a Learning Algorithm

- ✓ **Video:** Deciding What to Try Next  
5 min
- ✓ **Video:** Evaluating a Hypothesis  
7 min
- ✓ **Reading:** Evaluating a Hypothesis  
4 min
- ✓ **Video:** Model Selection and Train/Validation/Test Sets  
12 min
- ✓ **Reading:** Model Selection and Train/Validation/Test Sets  
3 min

Bias vs. Variance

- ✓ **Video:** Diagnosing Bias vs. Variance  
7 min
- ✓ **Reading:** Diagnosing Bias vs. Variance  
3 min
- ✓ **Video:** Regularization and Bias/Variance  
11 min
- ✓ **Reading:** Regularization and Bias/Variance  
3 min
- ✓ **Video:** Learning Curves  
11 min
- ✓ **Reading:** Learning Curves  
3 min
- ▶ **Video:** Deciding What to Do Next Revisited  
6 min
- 📖 **Reading:** Deciding What to do Next Revisited  
3 min

Review

- ✓ **Reading:** Lecture Slides  
10 min
- 📋 **Quiz:** Advice for Applying Machine Learning  
5 questions
- 🔗 **Programming Assignment:** Regularized Linear Regression and Bias/Variance  
3h

Building a Spam Classifier

- ▶ **Video:** Prioritizing What to Work On  
9 min
- 📖 **Reading:** Prioritizing What to Work On  
3 min

Learning Curves

Training an algorithm on a very few number of data points (such as 1, 2 or 3) will easily have 0 errors because we can always find a quadratic curve that touches exactly those number of points. Hence:

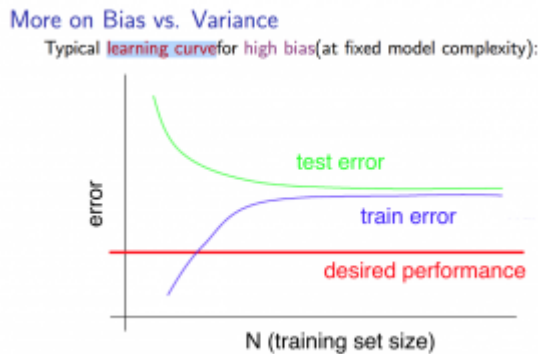
- As the training set gets larger, the error for a quadratic function increases.
- The error value will plateau out after a certain m, or training set size.

Experiencing high bias:

**Low training set size:** causes  $J_{train}(\Theta)$  to be low and  $J_{CV}(\Theta)$  to be high.

**Large training set size:** causes both  $J_{train}(\Theta)$  and  $J_{CV}(\Theta)$  to be high with  $J_{train}(\Theta) \approx J_{CV}(\Theta)$ .

If a learning algorithm is suffering from **high bias**, getting more training data will not **(by itself)** help much.

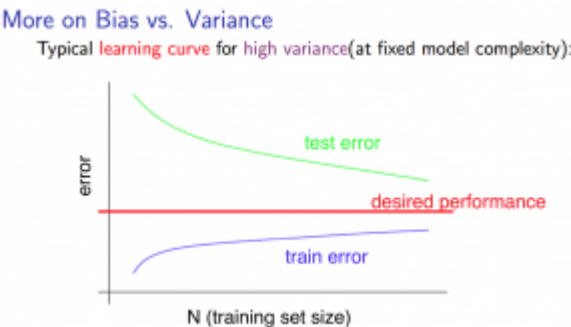


Experiencing high variance:

**Low training set size:**  $J_{train}(\Theta)$  will be low and  $J_{CV}(\Theta)$  will be high.

**Large training set size:**  $J_{train}(\Theta)$  increases with training set size and  $J_{CV}(\Theta)$  continues to decrease without leveling off. Also,  $J_{train}(\Theta) < J_{CV}(\Theta)$  but the difference between them remains significant.

If a learning algorithm is suffering from **high variance**, getting more training data is likely to help.



✓ Complete

Go to next item