

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

Diagnosing Bias vs. Variance

In this section we examine the relationship between the degree of the polynomial  $d$  and the underfitting or overfitting of our hypothesis.

- We need to distinguish whether **bias** or **variance** is the problem contributing to bad predictions.
- High bias is underfitting and high variance is overfitting. Ideally, we need to find a golden mean between these two.

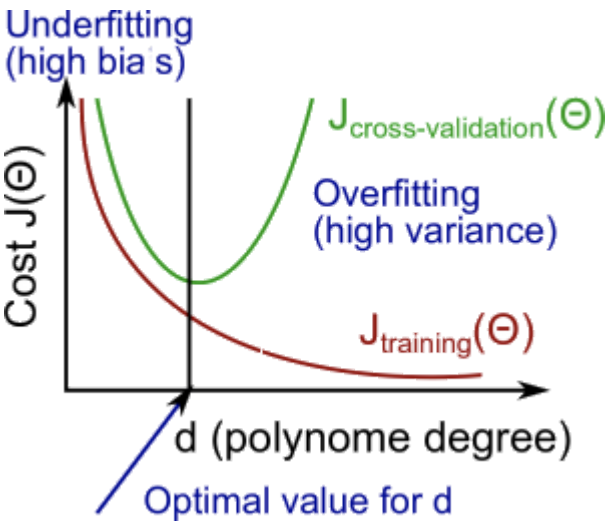
The training error will tend to **decrease** as we increase the degree  $d$  of the polynomial.

At the same time, the cross validation error will tend to **decrease** as we increase  $d$  up to a point, and then it will **increase** as  $d$  is increased, forming a convex curve.

**High bias (underfitting):** both  $J_{train}(\Theta)$  and  $J_{CV}(\Theta)$  will be high. Also,  $J_{CV}(\Theta) \approx J_{train}(\Theta)$ .

**High variance (overfitting):**  $J_{train}(\Theta)$  will be low and  $J_{CV}(\Theta)$  will be much greater than  $J_{train}(\Theta)$ .

The is summarized in the figure below:



✔ Complete

Go to next item