

## 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

## 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
-  **Video:** Error Analysis  
13 min
-  **Reading:** Error Analysis  
3 min

## Handling Skewed Data

- ▶ **Video:** Error Metrics for Skewed Classes  
11 min
- ▶ **Video:** Trading Off Precision and Recall  
14 min

## Using Large Data Sets

-  **Video:** Data For Machine Learning  
11 min

## Review

- ✓ **Reading:** Lecture Slides  
10 min
- ✓ **Quiz:** Machine Learning  
System Design  
5 questions

## Model Selection and Train/Validation/Test Sets

Just because a learning algorithm fits a training set well, that does not mean it is a good hypothesis. It could over fit and as a result your predictions on the test set would be poor. The error of your hypothesis as measured on the data set with which you trained the parameters will be lower than the error on any other data set.

Given many models with different polynomial degrees, we can use a systematic approach to identify the 'best' function. In order to choose the model of your hypothesis, you can test each degree of polynomial and look at the error result.

One way to break down our dataset into the three sets is:

- Training set: 60%
- Cross validation set: 20%
- Test set: 20%

We can now calculate three separate error values for the three different sets using the following method:

1. Optimize the parameters in  $\Theta$  using the training set for each polynomial degree.
2. Find the polynomial degree  $d$  with the least error using the cross validation set.
3. Estimate the generalization error using the test set with  $J_{\text{test}}(\Theta^{(d)})$ , ( $d$  = theta from polynomial with lower error);

This way, the degree of the polynomial  $d$  has not been trained using the test set.