

# **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

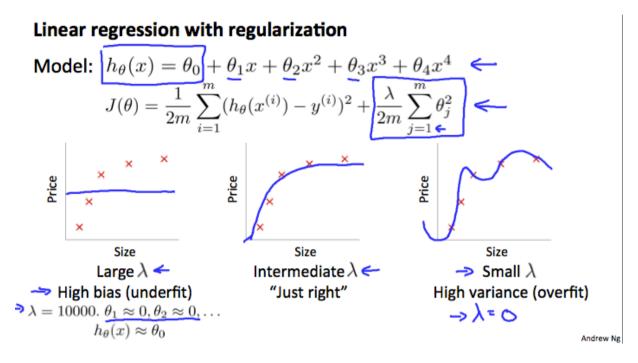
- Reading: Lecture Slides
  10 min
- Quiz: Advice for Applying
  Machine Learning
  5 questions
- Programming Assignment:
  Regularized Linear
  Regression and
  Bias/Variance
  3h

#### **Building a Spam Classifier**

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

## Regularization and Bias/Variance

**Note:** [The regularization term below and through out the video should be  $\frac{\lambda}{2m}\sum_{j=1}^n\theta_j^2$  and **NOT**  $\frac{\lambda}{2m}\sum_{j=1}^m\theta_j^2$ ]



In the figure above, we see that as  $\lambda$  increases, our fit becomes more rigid. On the other hand, as  $\lambda$  approaches 0, we tend to over overfit the data. So how do we choose our parameter  $\lambda$  to get it 'just right' ? In order to choose the model and the regularization term  $\lambda$ , we need to:

- 1. Create a list of lambdas (i.e. λ∈{0,0.01,0.02,0.04,0.08,0.16,0.32,0.64,1.28,2.56,5.12,10.24});
- 2. Create a set of models with different degrees or any other variants.
- 3. Iterate through the  $\lambda$ s and for each  $\lambda$  go through all the models to learn some  $\Theta$ .
- 4. Compute the cross validation error using the learned  $\Theta$  (computed with  $\lambda$ ) on the  $J_{CV}(\Theta)$  without regularization or  $\lambda = 0$ .
- 5. Select the best combo that produces the lowest error on the cross validation set.
- 6. Using the best combo  $\Theta$  and  $\lambda$ , apply it on  $J_{test}(\Theta)$  to see if it has a good generalization of the problem.





