

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

Building a Spam Classifier

Handling Skewed Data



Deciding What to Do Next Revisited

Our decision process can be broken down as follows:

- **Getting more training examples:** Fixes high variance
- **Trying smaller sets of features:** Fixes high variance
- **Adding features:** Fixes high bias
- **Adding polynomial features:** Fixes high bias
- **Decreasing λ :** Fixes high bias
- **Increasing λ :** Fixes high variance.

Diagnosing Neural Networks

- A neural network with fewer parameters is **prone to underfitting**. It is also **computationally cheaper**.
- A large neural network with more parameters is **prone to overfitting**. It is also **computationally expensive**. In this case you can use regularization (increase λ) to address the overfitting.

Using a single hidden layer is a good starting default. You can train your neural network on a number of hidden layers using your cross validation set. You can then select the one that performs best.

Model Complexity Effects:

- Lower-order polynomials (low model complexity) have high bias and low variance. In this case, the model fits poorly consistently.
- Higher-order polynomials (high model complexity) fit the training data extremely well and the test data extremely poorly. These have low bias on the training data, but very high variance.
- In reality, we would want to choose a model somewhere in between, that can generalize well but also fits the data reasonably well.