Evaluating a Learning Algorithm

Bias vs. Variance

- Diagnosing Bias vs. Variance — 7 min
- Diagnosing Bias vs. Variance — 3 min
- Regularization and Bias/Variance — 11 min
- Regularization and Bias/Variance — 3 min
- Learning Curves — 11 min
- Learning Curves — 3 min
- Deciding What to Do Next Revisited — 6 min
- Deciding What to do Next Revisited — 3 min

Review

Building a Spam Classifier

Handling Skewed Data

Using Large Data Sets

Review

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 $J_{train}(\Theta) \approx J_{CV}(\Theta)$.

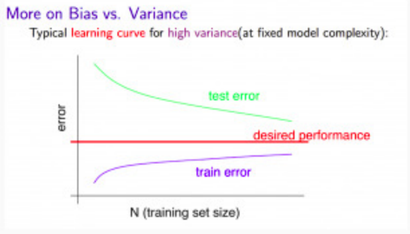
If a learning algorithm is suffering from **high bias**, getting more training data will not **(by itself)** help much.

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.



✓ Completado