

Advice for Applying ML

Repeatedly making good decisions can save a ton of time.

Debugging a learning alg.

- Get more training data
 - Smaller!
 - Additional features, polynomial features
 - increase/decrease learning rate λ
- There are some diagnostics that will advise.

Evaluating a model

- Plotting hard in $h > 2$ dims
- Compare J_{train} vs. J_{test} (regression) or count $y \neq \hat{y}$ (classification) to know whether ^{train vs. test} the model is generalizable

Model Selection and Cross-Validation

Note that when using J_{test} to compare models (e.g. poly degree), you're essentially adding another feature based on test data!

→ Split into three sets $\left\{ \begin{array}{l} \text{train} \\ \text{test} \\ \text{validate (cross-val) (dev set)} \end{array} \right.$

- J_{train} to fit model
- Eval model w/ J_{cv} , choose the best
- Estimate error w/ J_{test}

(Applies to NNS too, comparing architectures)

⚠️: don't use test data to build any model!

