Deciding What to Try Next

Evaluating a Learning Algorithm

Advice for Applying Machine Learning:

Introduction

- how to choose one of the most promising avenues to spend your time pursuing.
  - number of practical suggestions,
  - advice,
  - guidelines
- on how to do that.
- And concretely what we'd focus on is the problem of, suppose you are developing a machine learning system or trying to improve the performance of a machine learning system, how do you go about deciding what are the proxy avenues to try next?

Debugging a learning algorithm:

Suppose you have implemented regularized linear regression to predict housing prices.

$$J(\theta) = \frac{1}{2m}\left[\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2 + \lambda\sum_{j=1}^{m}\theta_j^2\right]$$

However, when you test your hypothesis on a new set of houses, you find that it makes unacceptably large errors in its predictions. What should you try next?

Actions

- Some actions
  - Get more training examples
  - Try smaller set of features (a small set of features)
  - Try getting additional features (just the opposite)
  - Try adding polynomical features
  - Try increasing lambda
  - Try decreasing lambda
- People generally randomly choose one and try it, which is waste of time most of the time.

Machine learning diagnostic:

Diagnostic: A test that you can run to gain insight what is/isn't working with a learning algorithm, and gain guidance as to how best to improve its performance.

Machine learning diagnostic:

Diagnostic: A test that you can run to gain insight what is/isn't working with a learning algorithm, and gain guidance as to how best to improve its performance.

Diagnostics can take time to implement, but doing so can be a very good use of your time.

Exercise

- Which of the following statements about diagnostics are true? Check all that apply.
  - It's hard to tell what will work to improve a learning algorithm, so the best approach is to go with gut feeling and just see what works.
  - Diagnostics can give guidance as to what might be more fruitful things to try to improve a learning algorithm.
  - Diagnostics can be time-consuming to implement and try, but they can still be a very good use of your time.
  - A diagnostic can sometimes rule out certain courses of action (changes to your learning algorithm) as being unlikely to improve its performance significantly.