Speeding up your ML code

Lyle Ungar

Photo credit http://allinguide.com/best-tips-how-to-speed-up-your-wordpress-website-or-blog/
Your ML code runs too slow; What can you do?
How to speed up your ML?

- **Speed up the code**
  - Use a faster language
  - Use a cluster/multicore machine /GPU
  - Vectorize

- **Use a streaming algorithm**
  - In features or observations

- **Develop on a subset of the data**
  - Or a subset of the features (univariate preprocessing)

- **Do dimensionality reduction**

- **Pick a faster algorithm**
  - Logistic regression ➔ ?
  - Kernelized SVM ➔ ?
  - Stepwise regression ➔ ?
  - K-NN ➔ ?
Pick a faster algorithm

• Logistic regression ➔ linear regression
• Kernelized SVM ➔ linear SVM
• Stepwise regression ➔ stagewise regression
• Knn ➔ K-means
How to speed up your ML: True/False

• Sparse code runs faster?
• Vector-based code runs faster?
• Models based on principle components are usually faster than one in the original features?