Chapter 9: Machine Learning

“The roots of statistics, as in science, lie in working with data and checking theory against data.”

“In the past fifteen years, the growth in algorithmic modeling applications and methodology has been rapid. It has occurred largely outside statistics in a new community—often called machine learning—that is mostly young computer scientists.”

Leo Breiman, Statistical Modeling: The Two Cultures

We will want to cover:

* model selection, cross validation, bootstrapping and train/test data
* prediction versus inference
* parametric versus not parametric
* supervised versus unsupervised

Trying to cover machine learning in one chapter is like a grown man trying to slide into a newborn’s onesie, some things just aren’t going to fit. We will talk through what it is, and so on. But there is a pretty big difference between applying machine learning and applying machine learning well.

* What’s possible with machine learning?
  + quantitative prediction and inference about variables
  + classification and clustering
* Side-bar (perhaps): parametric versus non-parametric

Recommended Reading