Lecture Notes 0 – Intro to Machine Learning

Marina Meilă
mmp@stat.washington.edu

Department of Statistics
University of Washington

September 30, 2021

What's in a name? Or where does "Machine Learning" /"Statistical Learning" come from?

- **What's in this sequence?**
  - Data analysis problems (e.g. clustering, classification)
  - Statistical models (e.g. exponential family models, graphical models)
  - Statistical methods (e.g. Support Vector Machines)
  - Algorithms (e.g. message passing, K-means). There is a continuum between algorithms, methods, and some of the other items on this list.
  - Mathematical facts/concepts from: graph theory, convex analysis
  - Theorems (without proofs), lemmas (with proofs)

Taxonomies

. . . all of them incomplete

- Statistical Learning Problems
  - Unsupervised
  - Supervised
  - (Semi-supervised)
  - Reinforcement
- Statistical models
  - Parametric
  - Non-parametric
- Statistical inference paradigms
  - Bayesian
  - Maximum Likelihood (ML)
  - Penalized Likelihood
  - Maximum A-Posteriori (MAP)

  These lists are meant to show that in this course we will not adopt a particular paradigm, but we will touch on most of them.

Plan for 535

- Supervised Learning (Prediction)
  - Predictor examples
  - Basic concepts: decision region, loss function, generative vs discriminative, bias-variance tradeoff
  - Training predictors: gradient descent, [Newton method]
  - [Combining predictors: bagging, boosting, additive models]
  - Regularized predictors: model selection, support vector machines, L1 regularization,
  - Learning theory and model selection basics
- Unsupervised Learning
  - Clustering: parametric, non-parametric
  - [Graphical models intro]
  - [Non-linear dimension reduction and geometric learning]
  - [Semi-supervised learning]

g graph data

- [Reinforcement Learning]