• 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.

• 535 “Unsupervised Learning” will have two major topics
  – Graphical probability models, i.e. algorithms for multivariate statistical inference
  – Unsupervised learning, mainly clustering

• 538 “Supervised Learning” will have two major topics
  – Supervised learning (classification, regression)
  – Convexity and convex optimization in statistics