Theory of Machine Learning

Spring 2022 — Lecture 1

Course logistics, basics

Basic logistics

- Class hours:
    - Tue/Th: 10:45 - 12:05
    - Instructor: Aditya Bhaskara (MEB 3470)
    - Email: a.bhaskara@utah.edu
    - TA: Chris Harker (chris.harker@utah.edu)
- Course webpage: Canvas

COVID Logistics

gcloud.utah.edu

- Can join via Zoom or in-person (throughout the semester)
- Recordings will be made available (mild delays possible)
- Lecture notes — scribe notes

Grading logistics

Lecture 2:

- Graduate class — (how to navigate)
- Four HWs — 60% of grade $(-\infty, +3]$ (± 3 days of deadline is ok
  Beyond that, ask...).
  - discussion encouraged — please write up on your own

Latex or Markdown.
↓
Pdf.

- Group project — 25% of grade (groups of 2 or 3 students)
  → [20 projects... reading + presenting + implementation].
- Scribe notes — 15% of grade
  (each student scribes one lecture — signup sheet + template)

(jot down notes during the lecture, polish after going back --)

Overview

- "Foundations" of modern machine learning
  - How are ML algorithms different? (need to generalize to inputs you have never seen...)
  - Optimization — what we can show and what we can't
  - Regularization generalization
  - Deep learning (some of the history, expressibility, hardness results, generalization, learning guarantees.)
  - Unsupervised learning (Bunch of data ⇝ discover underlying structure..)

SGD ; AdaGrad | Adam.

Accelerated GD.

Background in linear algebra, calculus, probability, ...

Four main themes

- Definitions, Valiant's PAC model

  Leslie Valiant.
  Probabilistically Approx Correct (PAC)

  - how to formally say, "algorithm works on unseen data"? (basic setup in ML.)
  - generalization bounds, VC dimension — why is minimizing training error (ERM) the right thing to do?! (regularization, etc.)
  - Rademacher complexity

- Optimization — core of any ML algorithm

  - convergence rates, how to choose learning rates
  - connection to "online" learning (+ boosting)

Four main themes

- Neural nets
  - basics, what they are and how difficult are they to train?
  - generalization, robustness, how to reason?

(much of the work is from 2015-on).

- Unsupervised learning
  - basic problems — clustering and representation learning
  - generative models