Review for Exam 2

Topics
These are the topics since Exam 1, although note that Exam 2 covers all topics in the course.

Generalization theory
- Inductive bias of choosing a function class
- Concentration of measure
- How to avoid over-fitting with finite function classes
- Rademacher complexity

You don’t need to understand the proof of the main Rademacher complexity result, nor the subsequent analysis that analyzes SVMs.

Convex optimization
- Convex sets and convex functions
- Tests for the convexity of functions
- Convex optimization problems (in standard form)
- The local-to-global property of convex optimization problems

Optimization algorithms
- Local optimization and gradient descent
- Role of step sizes
- Stochastic gradient method

You don’t need to understand the methods for constrained optimization.

Neural networks
- Using the chain rule in gradient computation
- Structure of and motivations for neural networks
- Forward and backward propagation algorithms

You don’t need to understand the matrix view of forward/backward propagation, nor any of the issues with “modern networks”.

Classification objectives
- Scoring functions
- Surrogate losses
- Cost-sensitive risks
- Importance-weighted risks
- Eliciting conditional probabilities with loss functions
- Multiclass classification
- One-against-all reduction

You don’t need to understand probabilistic calibration.

But you should the rest very well.

**Ensemble methods**

- Motivation for majority vote ensembles
- Decision trees
- Bagging, random forests
- AdaBoost

You don’t need to understand the theory behind AdaBoost.

For the exam, you should just know what these various methods are, and what they assume/exploit about the base predictors.

**Societal consequences**

- How machine learning can compromise privacy
- Concept of differential privacy
- How machine learning can be unfair
- Statistical parity, equalized odds
- Issues with the nature of training data

You don’t need to understand the details about how the Laplace mechanism works, nor the calibration/COMPAS issues.

**Clustering and PCA**

You will not be responsible for these topics for the exam.

**Big ideas**

- Statistical model for predictions and prediction functions
- Optimal predictions and optimal prediction functions
- The plug-in principle (e.g., model parameters, empirical distribution)
- Risk and empirical risk
- Decision boundaries
- Linear functions and feature expansions
- Inductive bias
- Concentration inequalities and generalization
- Aggregation/ensemble methods (e.g., model averaging, online-to-batch, bagging, AdaBoost)
- Mathematical and convex optimization as a tool to formulate learning methods
- Local optimization methods
- Computation graphs
- Cost-sensitive and importance-weighted risks
- Predicting conditional probabilities
- Reductions
- Societal consequences of machine learning