



Theoretical Foundations of Deep Learning: Course Overview

CS 839@UW-Madison
Yingyu Liang



Teaching team



- Yingyu Liang
email: yliang@cs.wisc.edu
office hours: Thu 1-2pm
office: CS building 5387
- Graders: Parashar Akshay, aparashar@wisc.edu
- More Information on course website:
https://pages.cs.wisc.edu/~yliang/cs839_spring22/index.html

Where to find content



- Canvas – private materials that should not be shared
 - Assignments
 - Grades
- Course website – public materials
 - Slides
 - Schedule
 - Policies
- Piazza
 - Discussion, questions
 - Announcements

Major goals



1. Learn about open questions/challenges in theoretical analysis of deep learning
2. Learn about existing techniques/results
3. Try to propose new questions/techniques and get new results

Emphasize on **open discussion**

May not fit: if just interested in **using** deep learning or the empirical aspects, but not about its theoretical analysis

Discussions



1. What is **Deep Learning**?
2. What is the meaning of **Theoretical Foundations**?
3. Why care about theoretical analysis?

Course requirements



- ~5 homework assignments: 40%
 - Math analysis
 - **post on Canvas; submit your solutions on Canvas**
- Paper presentation: 10%
- Lecture note scribe: 10%
- Project: 40%

Expected background

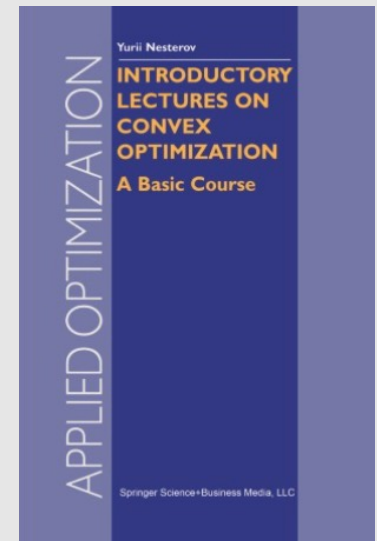
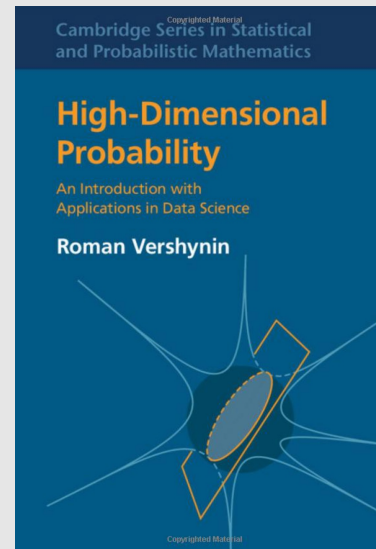
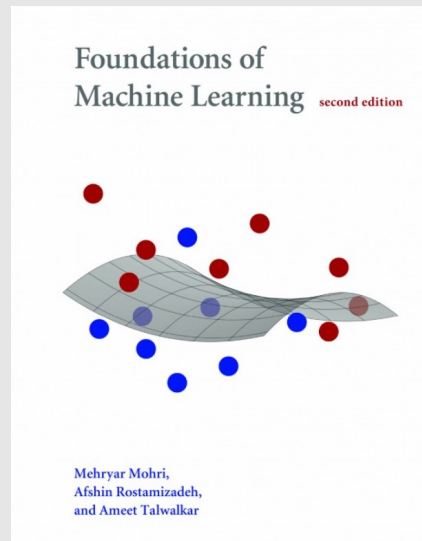
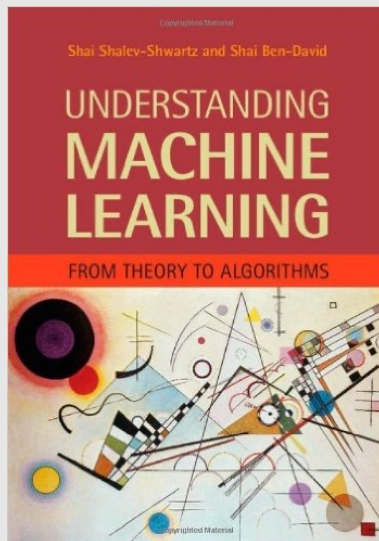


- (CS760 or CS761 or CS861)
- **AND** strong math background in
 - machine learning
 - probability/statistics
 - optimization

Standard Tools



- Understanding machine learning: From theory to algorithms. Shai Shalev-Shwartz, and Shai Ben-David. Cambridge University Press, 2014.
- Foundations of machine learning. Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. MIT Press, 2018.
- High-Dimensional Probability: An Introduction with Applications in Data Science. Roman Vershynin. Cambridge University Press, 2018
- Introductory Lectures on Convex Optimization: A Basic Course. Yurii Nesterov. Springer, 2004.



Topics



- Challenges in theoretical analysis of deep learning
- Approximation/representation power of neural networks
- Neural Tangent Kernel (NTK)
- Mean-field analysis
- Optimization: landscape, Implicit regularization
- May consider: complexity, representation learning, GAN, ...



Assignments



Review reading



- Understanding machine learning: From theory to algorithms. First 4 chapters.
- High-Dimensional Probability: An Introduction with Applications in Data Science. Chapter 1, 2, 8
- Links can be found on course website:
https://pages.cs.wisc.edu/~yliang/cs839_spring22/index.html

HW1: Background test



- Will post on Canvas
- Will set up how to submit the solutions on Canvas
- If you find many questions intimidating, we suggest you drop the course and take it again in the future when you are more prepared

Topics:

- ML theory basics
- probability
- Optimization basics

THANK YOU

