Foundations of Deep Learning, Spring 2022

CS-GY 9223K / ECE-GY 9133 A

Objectives

The impact of deep learning in numerous areas of science, engineering, and technology has never been higher than it is today. However, progress in practical applications of deep learning has considerably outpaced our understanding of their foundations. Remarkably, many fundamental questions remain unanswered. Why are we able to train them so efficiently at all? Why do they perform so well on unseen data? Is there any benefit of one neural architecture over another?

The goal of this course is to:

- outline the nascent body of work on the theoretical foundations of deep learning, and
- introduce students to a variety of tools from approximation theory, optimization, and statistical learning theory that may help inform our understanding of deep neural networks,
- spur students to pursue research in this exciting new area of machine learning.

Caution: This course is designed for advanced graduate students. This will entirely be a pen-and-paper, seminar-style course. Students interested in practical aspects of deep learning should instead consider taking CS-GY 6953, ECE-GY 7123, or equivalent.

Pre-requisites

- CS-GY 6953, ECE-GY 7123, or equivalent graduate-level course on Deep Learning
- Mathematical maturity (in particular: probability, linear algebra and optimization).
- An appetite for reading theory-style papers.

Outline of lectures (tentative)

Week	Lecture topic
1	Representations: Universal approximation
2	Representations: The impact of depth
3	Learning: Local convergence
4	Learning: Shallow models
5	Learning: The NTK regime
6	Learning: Beyond NTK
7	Generalization: Implicit regularization
8	Generalization: PAC learning and uniform stability
9	Generalization: Double descent
10	Misc topics: Certified robustness
11	Misc topics: Differentiable games
12	Misc topics: Privacy and security
13	Misc topics: Reinforcement Learning
14	Wrap-up

Grading policy

We will not have homework assignments or exams. Instead, evaluation will be performed according to a (semester long) project. Projects will be in groups of at most 2, and will involve you surveying a specific topic involving theoretical aspects of deep learning. Here, a "topic" is defined as any line of inquiry with 2 or more papers. Deliverables include:

- A project proposal due one month into the course identifying the specific topic (and list of papers).
- A 10-minute (video) presentation due two months into the course summarizing progress done so far.
- A final report, along with any accompanying code.
- A 20-minute (video) presentation detailing your report.

Contact info

Chinmay Hegde (chinmay.h@nyu.edu)

Resources

There is no textbook for this course. Instead, our primary reading material will the course notes posted during each lecture week. Useful external resources include:

- Fit without Fear by Misha Belkin.
- Mathematics of Deep Learning by Joan Bruna.
- Science of Deep Learning by Aleks Madry and Costis Daskalakis.
- Theoretical Foundations of Deep Learning by Ankur Moitra.
- Deep Learning Theory by Matus Telgarsky.

Course times

Lectures: Mondays 11am-1:30pm.

Instructor office hours: By appointment.

Inclusivity

It is my intent that *all* students' learning experiences are rewarding both in and out of class. If this standard is not being upheld, please speak to me. I cherish your feedback. The Diversity and Inclusion at Tandon website is a terrific resource. If you are a student with a disability who is requesting accommodations, please reach out NYU's Moses Center for Students with Disabilities at 212-998-4980 or mosescsd@nyu.edu.