[Science of Deep Learning] Syllabus

Instructor Information

Instructor	Email	Office Location & Hours
Ziming Liu	zmliu@mit.edu	TBA

General Information

Description

Recent advancements in deep learning are mostly driven by engineering rather than science. While engineering is very effective in making progress in the short run, science plays a complementary role by providing lasting insights into understanding and advancing deep learning models. Many existing courses on machine learning theory focus on classic statistical theories, which are mathematically rigorous but may not be suitable to explain rich phenomena in modern deep learning, e.g., neural scaling laws and emergent abilities. This course takes a physicist's approach to understanding, emphasizing the role of phenomenology (intuitive mental pictures) in bridging the gap between theories and experiments. Since deep neural networks are complex systems that might be too complicated to admit a unified theory (yet), the course aims to cover views of deep neural networks as broadly as possible: Neural networks as feature learners, information systems, geometry/manifold learners, dynamical systems, complex systems, biological systems, etc. These views will hopefully give students versatile aspects of deep neural networks and help them develop their own views and tastes in research. I will also demonstrate the "hypothesis-driven" research style (instead of "performance-driven"), which is commonly adopted in scientific research but largely glossed over in deep learning research.

Prerequisites

Math: Basic level of familiarity with linear algebra and calculus. Familiarity with more advanced math topics (differential equations, differential geometry) is preferred but not mandatory (we will cover them when necessary).

Programming: Familiar with Python (or any programming language like Jax or Julia that conveniently supports training deep neural networks). Having experience in training neural networks is preferred.

Expectations and Goals

I hope that students who come from either theoretical backgrounds or engineering backgrounds find this course approachable and learn things that are complementary to their existing skill sets. At the end of this course, (1) Students can see the global and systematic picture of deep learning. (2) Students understand the difference between scientific and engineering understanding and can flexibly apply both to their research.

Course Schedule (Tentative)

Week	Topic
[Week 1 (Intro)]	Introduction: Philosophy of “Science of deep learning”, Introduction to DL models and applications
[Week 2 (Classic)]	Statistical Learning Theory
[Week 3 (Classic)]	Statistical Physics Approaches
[Week 4 (Classic)]	Optimization Theory
[Week 5 (Modern)]	Neural networks as feature learners: initialization, NTK, MuP, initial condensation, random feature models
[Week 6 (Modern)]	Neural networks as information systems: information bottleneck, language modeling as compression, memory networks, Hopfield networks, information interpretation of transformers
[Week 7 (Modern)]	Neural networks as geometry/manifold learners: manifold learning methods, diffusion models, equivariant neural networks
[Week 8 (Modern)]	Neural networks as dynamical systems I (dynamics of hidden states): Recurrent neural networks, Neural ODE, transformers & in-context learning
[Week 9 (Modern)]	Neural networks as dynamical systems II (dynamics of optimization): Langevin dynamics for SGD, Edge of Stability, mode connectivity, Repon model
[Week 10 (Modern)]	Neural networks as complex systems: emergent behaviors and phase transitions, double descent, grokking, neural scaling laws
[Week 11 (Modern)]	Neural networks as biology: Mechanistic interpretability
[Week 12 (Next-gen)]	Math for AI: Kolmogorov-Arnold Networks (KAN), etc
[Week 13 (Next-gen)]	Physics for AI: Boltzmann machines, physics-inspired generative models
[Week 14 (Next-gen)]	Biology/Neuroscience for AI: Hebbian learning, spiking neural networks
[Week 15 (Applications)]	AI for Science applications
[Week 16 (End)]	Final exam or course project presentations