APMTH 226: Neural Computation - Fall 2024

03/25/2024
This document is subject to change

Instructor: Prof. Cengiz Pehlevan, cpehlevan@seas.harvard.edu,

Office hours will be announced on course website.

Teaching fellow: William Tong, wtong@g.harvard.edu

TBD

Description: This course is an introduction to the theory of computation with neural networks. We will cover selected topics from theoretical neuroscience and deep learning theory with an emphasis on topics at the research frontier. These topics include expressivity and generalization in deep learning models; infinite-width limit of neural networks and kernel machines; deep learning dynamics; biologically-plausible training of neural networks and models of synaptic plasticity; reinforcement learning in the brain; neural population codes; normative theories of sensory representations; computing with dynamics in recurrent neural networks; attractor network models of memory and spatial maps; probabilistic generative modeling and diffusion models; sequence modeling and transformers

Course goal and prerequisites: This course is intended for students starting research in deep learning theory, theoretical neuroscience and theory of neural computation. The aim is to introduce some of the basic mathematical techniques and models used in the field. At the end of the course, you will be comfortable with reading research papers in the field. Mathematical maturity and programming skills are important for benefitting from the course. You should be able to solve the prerequisite problems that will be posted on the canvas site.

Course Structure:

1. We will have two 75 minute live lectures per week. Lecture times: M/W 9:00-10.15am. Location: TBD

2. There will a one 90 minute TF session per week. We may have more depending on extra TF support we get. Time will be announced. There won't be any lectures on this hour, TFs will answer questions on the course material, homework and final projects.

Grading: 70% Problem Sets, 30% Final Project

Problem sets and final project

Biweekly problem sets: Problem sets will typically include a coding problem and a few analytical problems. Late submissions will incur 10 points off per day late unless permission from the instructor. Permission will be granted for exceptional case. No homework is accepted once solutions are posted (which may be anytime after the deadline).

Final project: A typical project will involve reproducing the results of a research paper and extending it.

References: We will not follow a particular textbook, but the following are useful:

- Theoretical neuroscience – Dayan and Abbott
- Introduction to the Theory of Neural Computation – Hertz, Korgh and Palmer
- Theory of Neural Information Processing – Coolen, Kuln and Sollich
- Mathematical Foundations of Neuroscience – Ermentrout

Most of course material will be drawn from research papers. A list will be made available in Canvas.

Lecture notes will be provided.

Topics: Selected topics from below. We will cover about 75%.

<u>Module 1- Learning</u>

1. Basics of learning from examples, Perceptron learning algorithm and the convergence theorem, Perceptron capacity and Cover's counting theorem, Gardner's statistical mechanical analysis of perceptron capacity, basics of Statistical Learning Theory
2. Expressivity and generalization in deep learning models
3. Infinite-width limit of deep learning and kernel machines, the Neural Tangent Kernel, random feature models
4. Feature learning infinite-width limits and dynamical mean field theory
5. Deep networks and the brain, The backpropagation algorithm and biologically-plausible alternatives to backpropagation

6. Representation learning in the brain: Hebbian plasticity and unsupervised learning, Linear networks and PCA, Nonlinear networks and ICA, disentangled representations
7. Probabilistic generative modeling: Diffusion models
8. Reinforcement learning: Basics and its implementation in the brain

<u>Module 2- Neural Codes: Efficiency and Noise</u>

1. Infomax and retinal ganglion cells
2. Sparse coding and early visual representations
3. Noisy population codes: basics of statistical estimation, Fisher information, Cramer-Rao Bound, discriminability
4. Noise correlations and effect on coding

<u>Module 3- Dynamics</u>

1. Single neuron models: Hodgking-Huxley, Morris-Lecar, Integrate-and-fire neuron, noisy IF-neuron as a threshold-crossing process
2. Network dynamics: Computing with transients. Dynamical mean field theory of Randomly Connected Networks and Reservoir Computing
3. Network dynamics: Computing with attractors. Hopfield networks and associative memory, Lyapunov stability, Attractor models of short term memory, Attractors and the brain's navigation system
4. Training recurrent neural networks: Backpropagation Through Time, Boltzmann Machines
5. Sequence-to-sequence models: Attention mechanism and transformers

Academic Integrity Policy:

Discussion and the exchange of ideas are essential to doing academic work. For assignments in this course, you are encouraged to consult with your classmates as you work on problem sets. However, after discussions with peers (or course instructional staff such as tutors, TF/TAs, course assistants), make sure that you can work through the problem yourself and ensure that any answers you submit for evaluation are the result of your own efforts. In addition, you must cite any books, articles, websites, lectures, etc that have helped you with your work using appropriate citation practices. Similarly, you must list the names of students with whom you have collaborated on problem sets.