Basic Information

Course: EE660: Mathematical Foundations and Methods of Machine Learning
Term: Spring 2024
Instructor: Stephen Tu (stephen.tu@usc.edu)
TAs:
Fernando Valladares Monteiro (fvallada@usc.edu)
Mutian Zhu (mutianzh@usc.edu)

Please include [EE660] in the subject of any email related to the course.

Time and Location:

Lecture
MW 4-5:50pm, OHE 136.

Discussion
F 11-11:50am, OHE 136.

Final:
May 1, 4:30-6:30pm, Location TBD.

Office Hours:
Tu 9-10am (Mutian), PHE320.
W 3-4pm (Stephen), EEB 326.
F 9:30-10:30am (Fernando), EEB322.

Course Prereqs: EE 503, EE 510, and EE 559

Course Description

This course will cover the mathematical foundations of machine learning, including both supervised/unsupervised learning, in addition to the basic principles of sequential decision making. The focus will be on the mathematical principles underlying modern machine learning algorithms. Topics covered include (but are not limited to) concentration inequalities, empirical process theory, margin theory, stochastic gradient methods, algorithmic stability, PAC-Bayes inequalities, generative modeling (energy-based models, variational autoencoders, diffusion models), optimal control, Markov Decision Processes, linear quadratic regulators, Kalman filtering, multi-arm bandits, contextual bandits, and basic reinforcement learning algorithms. Students will leave this course with a rigorous understanding of modern machine learning methods.

Textbook

There is no textbook for the course. Instead the reading will be from the course lecture notes.

Note: I will frequently update the lecture notes as the course progresses (filling in new content as we move forward in addition to correcting typos). While I will update the DEN with newer versions of the notes, I recommend instead using the link (https://stephentu.github.io/pdfs/EE660_Lecture_Notes.pdf) and refreshing each time you go to read.

Furthermore, as these notes are a work in progress, please feel free to send me an email if you suspect any typos or errors.

Homeworks

There will be 5 homeworks assignments throughout the semester. These homeworks will mostly be pen and paper exercises, with an occasional coding exercise. For any coding exercises, you are free to use whatever language you choose (although we will highly encourage the use of Python and its comprehensive scientific computing ecosystem).

The main deliverable will be a PDF file containing your solutions. **It is heavily encouraged that you use LaTeX to typeset your homework (a sample LaTeX template will be provided before the first homework is assigned)!**

Collaboration Policy: You are free to collaborate with other students on the homework assignments. **However, all solutions must be written by yourself.**

Late Policy: Late submissions will not be accepted, no exceptions! However, we will automatically drop your lowest score.

Grading breakdown

The following table describes how your final grade will be calculated.

Assignment	**Weight**
Homeworks (5 homeworks in total, we will automatically drop the lowest score).	30%
Midterm	30%
Final	40%

Topics

Supervised Learning

- Perceptron [2]
    - Mistake bound.
    - Generalization bound.
- Empirical Risk Minimization (short introduction to empirical process theory) [6]
    - Basic concentration inequalities.
    - Uniform convergence and Rademacher complexity.
    - Loss functions and margin theory.
    - Chaining (Dudley's inequality).
- Alternative Perspectives on Empirical Risk Minimization [2]
    - Algorithmic Stability.
    - PAC-Bayes Inequalities.
- The Curiosities of Overparameterized Models [1]
    - Bias/variance trade-off. Double descent.
    - Optimization is made easier.
    - Which minima generalize?

Probabilistic Modeling

- Energy based models [2]
    - Max likelihood training.
    - Langevin dynamics.
    - Score matching.
    - Denoising score matching.
- Normalizing flows [1]
    - Finite compositions.
    - The log probability ODE.
- Variational Autoencoders and GANs [1]
    - Evidence lower bound (ELBO).
    - The GAN min/max loss and its behavior.
- Diffusion Models [2]
    - The DDPM variational inference approach.
    - The continuous SDE approach.

Sequential Decision Making

- Dynamical Systems, Markov Decision Processes, and Optimal Control [1]
    - Q-functions.
    - Bellman's equations.
- Linear dynamical systems [2]
    - Linear Quadratic Regulator.

    - Kalman Filtering.
- Bandits [4]
    - Multi-arm bandits.
    - Explore-then-commit.
    - The Upper Confidence Bound (UCB) algorithm.
    - Contextual bandits and LinUCB.
- Reinforcement learning [2]
    - Value/Policy iteration, policy gradients, and derivative-free methods.
    - Model-based tabular RL.

Schedule

Note: The syllabus will be updated every week with assigned readings corresponding to each lecture. The section(s) in the **Reading and HW** column refer to the lecture notes.

Date	Topics	Reading and HW
1/8	Intro, binary classification, Perceptron mistake bound.	Sections 1.1 and 1.2.1.
1/10	Perceptron risk bound, empirical risk minimization framework.	Sections 1.2.2, 1.3.1, and 1.3.2.
1/15	**Holiday: MLK's Birthday**	
1/17	Sub-Gaussian random variables, Laplace transform method, basic concentration inequalities.	Section B. **HW1 assigned.**
1/22	Bounds for finite hypothesis classes.	
1/24	Hinge and logistic losses, margin theory, Rademacher complexity.	
1/29	Rademacher symmetrization and contraction.	
1/31	Generalization bounds and explicit computations.	**HW1 due.**
2/5	Chaining and Dudley's inequality.	**HW2 assigned.**
2/7	Algorithmic stability framework, stability of SGD and Gibbs-ERM.	
2/12	PAC-Bayes inequalities and basic compression bounds.	

2/14	Curiosities of modern deep networks.	
2/19	**Holiday: President's Day**	
2/21	Energy based models, max-likelihood estimation, Langevin dynamics.	**HW2 due, HW3 assigned.**
2/26	Score matching and denoising score matching.	
2/28	Normalizing flows, log probability ODE, neural ODEs.	
3/4	Variational autoencoders, evidence lower bound, generative adversarial networks.	**HW3 due.**
3/6	**Midterm (Location TBD)**	
3/11	**Holiday: Spring Recess**	
3/13	**Holiday: Spring Recess**	
3/18	Denoising diffusion probabilistic models.	**HW4 assigned.**
3/20	Diffusion via stochastic differential equations.	
3/25	Dynamical systems, Markov Decision Processes, optimal control formulation, Bellman's equations.	
3/27	Linear Quadratic Regulator.	
4/1	Kalman Filtering.	**HW4 due.**
4/3	Multi-arm bandits (MAB), regret analysis, and explore-then-commit.	**HW5 assigned.**
4/8	The UCB algorithm for MAB.	
4/10	Contextual bandits, LinUCB	
4/15	Finish LinUCB analysis.	
4/17	Value and policy iteration, contractivity of Bellman's operator, policy gradients.	**HW5 due.**
4/22	Model-based RL.	
4/24	TBD.	