Machine Learning Via Sequential Interaction (MLSI)

Lecturer: Shuchin Aeron
TA: TBD Course Information

---

1 Basic Information

- **Instructor**: Shuchin Aeron, Halligan 144, shuchin@ece.tufts.edu
- **Main Textbook**: Bandit Algorithms. Authors: Tor Lattimore, Csaba Szepesvari. PDF available at: `https://tor-lattimore.com/downloads/book/book.pdf`
- **Teaching Assistants/Co-Instructor**: TBD
- **Office Hours**: TBD
- **Course Website**: The course will be actively managed via Piazza as well as using Gradescope. You will receive an email to enroll in the course on Piazza. All the exams and hw will be managed via gradescope and all the class discussions, class notes, will be managed via Piazza.

1.1 Other resources

1. A Modern Introduction to Online Learning. Author: Francesco Orabona. `https://arxiv.org/abs/1912.13213`
2. Prediction, Learning, and Games. Authors: Nicola Cesa-Bianci, Gabor Lugosi.
3. Foundations of Reinforcement Learning and Interactive Decision Making. Authors: Dylan J Foster, Alexander Rakhlin, `https://arxiv.org/abs/2312.16730`.

2 Overview

This course will introduce machine learning as a sequential and interactive game (also technically referred to as online learning) in which the learner aims to minimize the regret: loss incurred for not playing the best strategy (in hindsight) averaged over time. This framework generalizes the traditional statistical learning set-up in a non-trivial manner to cases where the examples are not independent, are supplied only sequentially, and possibly in an adversarial manner. Topics include: general set-up of sequential learning (as bandit problems) with examples, multi-armed bandits, contextual bandits, practical algorithms and their analysis with provable guarantees, online convex and online linear optimization, online-to-batch conversion. Connections with and extensions to reinforcement learning will be highlighted with some contemporary applications.

3 Tentative schedule and coverage from the main textbook

Note: The HW are issued at points when we are done covering the necessary material as indicated below. The due date for each HW will also depend on the length of the assignment and exam schedule.

1. Lecture 1 - Course Outline. Chapter 1: Introduction to bandit problems.
2. Lecture 2 - Chapter 2: Foundations of Probability.

3. Lectures 3: Chapter 4: Stochastic Bandits. Issue HW 1.
4. Lecture 4, 5 - Chapter 5: Necessary results from concentration of measure.
5. Lecture 6 - Chapter 6: Explore Then Commit (ETC) algorithm. Issue HW 2.
6. Lecture 7, 8. - Chapter 7: The Upper Confidence Bound (UCB) algorithm.

   Mid-term I: Take-home exam.
7. Lecture 10, 11 - Chapter 11: Adversarial Bandits, The Exp3 algorithm.
8. Lecture 12 - Chapter 12: The Exp3-IX algorithm. Issue HW 3.
9. Lecture 13, 14 - Chapter 18: Contextual Bandits.
10. Lecture 15,16 - Chapter 19: Stochastic Linear Bandits. Issue HW 4.

    Mid-term II: Take-home exam.
11. Lecture 17, 18 - Chapter 26: Foundations of Convex Analysis.
12. Lecture 19, 20 - Chapter 27: Exp3 for Adversarial Linear Bandits.
13. Lecture 21, 22 - Chapter 28: Follow the Regularized Leader, Mirror Descent. Issue HW 5.
14. Lecture 23 - end - Dueling bandits, RLHF, Other topics.

Final take-home exam.

3.1 Homework and Exams

There will be 5 HW, 2 midterm exams, and a final Exam. The weight distribution for the grade is as follows.

1. **HW** - 50% **NOTE:** The HW are very important in this course. Please do spend a lot of time solving for them and also go over the solutions that are provided by the instructor and TA.

   HW submission: All HW are to be submitted and graded via gradescope.

   HW delays: Delays are not permitted unless there is a strong reason. Please contact the instructor and TA beforehand if the delay is to be anticipated and each case will be dealt on an individual basis.

2. **Take-Home Mid-Terms** - 15% each totaling to 30%

3. **Take-Home Final Exam** - 20%