Overview: This course is a graduate level class on online learning. Specifically, we will study multi-armed bandits in stochastic and adversarial settings. Bandit problems are present in settings where decision-making is done in conjunction with learning and collecting new data. Such systems have an inherent tension between exploitation (using previously collected information to make the "best" decisions now) and exploration (making decisions that are not necessarily optimal, but better explores the search space to collect fresh data). Applications abound in a variety of settings, including online advertising, A/B testing, drug testing, recommendation systems, online resource allocation, reinforcement learning in games (e.g. AlphaGo Zero) and autonomous driving.

Pre-requisites: EE 381J (graduate level Probability, Statistics and Random Processes) or equivalent.

Textbook: 1. *Bandit Algorithms,* T. Lattimore and C. Szepesvari, To be published by Cambridge University Press, 2019. A preprint of the book is available on the web at: `https://tor-lattimore.com/downloads/book/book.pdf`

Course Policy: Attendance is expected. You are responsible for material covered in the reading assignments (even if not covered in class) as well as material covered in class that is not in the course notes/book. Homework will be assigned roughly every week. Late homeworks will normally not be accepted.

You may discuss homework problems with other students and are permitted to work in groups, but you are not allowed to simply copy from others. University disciplinary procedures will be invoked if any form of cheating is detected. Course and instructor evaluations will occur the last week of class. Academic adjustment for students with disabilities is made through the SSD Office:
`http://ddce.utexas.edu/disability/`

Online Platforms: Homework and related class material will be posted on Canvas – UT's course management platform: `http://canvas.utexas.edu`. We will be using
Piazza: `https://piazza.com/utexas/fall2018/ee381j` – an online discussion platform – for posting and answering questions regarding class and homework, and using *Gradescope* for electronic submission and grading of homework: `https://gradescope.com/courses/21080` You will need to create Gradescope account (you will receive an email regarding this) and details on the (pdf) scanning and uploading the homework.

Grading:

  (i) Homework: 30%

 (ii) Midterm Exam: 30%

(iii) Final Project: 40%

Syllabus Outline

1. Stochastic bandits with finite number of arms: Explore and commit algorithm, UCB algorithm and regret analysis

2. Adversarial bandits with finite number of arms: Exponential weighting and importance sampling, Exp3 algorithm and variants

3. Multi-armed Bandit (MAB) lower bounds: minimax bounds, problem-dependent bounds

4. Contextual bandits: Bandits with experts – the Exp4 algorithm, stochastic linear bandits, UCB algorithm with confidence balls (LinUCB and variants)

5. Contextual bandits in the adversarial setting: Online linear optimization, Follow The Leader (FTL) and Follow the Regularized Leader (FTRL), Mirror Descent

6. Combinatorial bandits, Bandits for pure exploration, Infinite-armed bandits and Monte-Carlo Tree Search

7. Bandits in a Bayesian setting: Bayesian learning and regret, Thompson sampling