DSO 699: Special Topics in Data Sciences and Operations
"Bandit Algorithms and Reinforcement Learning"

USCMarshall
School of Business

Syllabus- Spring 2025- 3 Units "Will meet 3 hours per week"
Professor: Adel Javanmard
Office: Bridge Hall 300 A
Email: ajavanma@usc.edu
Office hours: *By appointment only*

Course Description

Sequential decision-making is a cornerstone of modern artificial intelligence and machine learning, driving advancements in various fields. Bandit algorithms and reinforcement learning are at the heart of this dynamic area, enabling systems to make intelligent choices in uncertain environments by balancing exploration and exploitation. These techniques are crucial in a myriad of applications domains, ranging from personalized recommendations and dynamic pricing strategies to self-driving cars. Understanding these methods unlocks opportunities for groundbreaking research and practical solutions, making them indispensable for aspiring AI and machine learning professionals.

This Ph.D. level course provides an advanced introduction to the fundamental methods in Bandit Algorithms and Reinforcement Learning, which underpin these real-time decision-making systems. The first half of the course focuses on classical bandit algorithms, exploring problems where the balance between exploration and exploitation is key, and the system dynamics are well-understood. We will cover the bandit formulation, Bernoulli bandits, Thompson sampling, UCB, greedy algorithms, regret analysis, and Eluder dimension.

Building on this foundation, the second half of the course introduces reinforcement learning and strategic exploration techniques. These methods address dynamic optimization problems where state spaces are vast and system dynamics are unknown. Students will learn about Markov Decision Processes, Bellman Operator and Value Iteration, Policy Iteration, Q-learning, model-based reinforcement learning, and strategic exploration in RL, including the RMAX algorithm. In these settings, we often need to learn about the system while simultaneously making optimal decisions.

Throughout the course, we emphasize rigorous theoretical foundations and algorithmic development, illustrating key concepts through practical applications. Our goal is to equip doctoral students with the expertise and skills needed to conduct their own pioneering research in this dynamic field.

Learning Objectives

Bandit algorithms and Reinforcement learning are the two main modules that we cover in this course. Below, we list the learning objective under each of these modules.

Fundamentals of Bandit Algorithms:
Upon successful completion of this course, students will be able to:
o   Describe the multi-armed bandit problem and its significance in decision-making.
o   Implement and analyze various bandit algorithms, including Thompson Sampling, UCB, and greedy algorithms.
o   Conduct regret analysis to evaluate the performance of bandit algorithms.
o   Apply the concept of Eluder dimension to assess the complexity of bandit problems.

- Describe adversarial bandit environments, importance weighted estimators, implement the EXP3 algorithm and conduct its regret analysis
- Describe contextual bandits, linear bandits, bandits with Expert advice, EXP4 algorithm and its regret analysis

Reinforcement Learning Concepts:
Upon successful completion of this course, students will be able to:
- Describe the basics of Markov Decision Processes (MDPs) and their role in reinforcement learning.
- Implement Bellman Operator and Value Iteration techniques to solve MDPs.
- Utilize Policy Iteration and Q-learning to find optimal policies.
- Explore model-based reinforcement learning approaches for efficient decision-making.
- Develop strategic exploration techniques, including the RMAX algorithm, for navigating unknown environments.

Prerequisites and/or Recommended Preparation:

The goal of this course is to provide students with a rigorous foundations in algorithmic development and theoretical understanding of sequential decision making. Given that it is a Ph.D. level course, students are expected to be familiar with probability and linear algebra. Most importantly, the course language is **"proof language".** Students are expected to have substantive mathematical maturity, ability to write mathematical proof and do numerical experiments with programming in Matlab, Python or R.

Required Materials (Textbooks)

The following references are used for background readings:
R1: Tor Lattimore and Csaba Szepesvári. "*Bandit algorithms".* Cambridge University Press, 2020.
R2: Alekh Agarwal, Nan Jiang, Sham M. Kakade. *"Reinforcement learning: Theory and algorithms."* CS Dept., UW Seattle, Seattle, WA, USA, Tech. Rep (2019).

I will also post a list of papers for the final projects on Brightspace.

Course Notes:

The majority of the course will be lecture based. A rough list of topics and outline of material is given in the "course description" and the "course calendar", however, depending on the speed of the class and discussion, these topics are open to change. The precise dates of homework and exams, however, will **not** change without substantive notice.

Grading Policies:

Your final course grade will be assessed as follows:

Assignment	% of Total Grade
Participation/Discussion	5%
Homework	30%
Midterm Exam	35%
Final Project/Presentation/Report	30%
Total	100%

Class Attendance & Participation

Students are expected to attend all lectures and more importantly actively contribute to lectures and discussions. Your comments and feedbacks on the course materials, and the pace of the course, or my teaching style are very welcome especially that this is the first time the course is offered.

If you are not comfortable bringing up your questions in the class, take advantage of the many opportunities to speak with me one-on-one. I am always accessible by e-mail, and will be more than happy to speak with you before or after class. Note: if your question requires a conversation rather than a short answer, email is not the best way to go – please talk to me after class or schedule an appointment with me to talk in office hours.

Homework Assignments

There will be three homework assignments (each 10%). Assignments should be written in Latex with a copy of the code included (if applicable). Assignments should be emailed to the instructor before 11:59 pm on the due date. Late homework are not acceptable.

Makeup Exam:

The midterm exam will be in-class based on the martials covered in the first half of the course. There will be no-make up for the midterm exam.

Final Projects

In the second half of the course, I will release a pool of references (recent papers on topics we cover in the course). Students should select one of these references (each reference can be chosen by only one student). Students are required to make a presentation to discuss the paper, and also submit a **5 page latex document** summarizing the paper. The reports should in particular discuss the main idea of at least one of the theorems in the paper.
The final projects are evaluated based on the quality of the report and the presentation, and the student's grasp of the material and ability to answer questions during the presentation. Final reports should be emailed to the instructor before 11:59 pm on the due date. Late reports are not acceptable.

Policy on Group Work

Discussion of homework assignments is permitted and encouraged; however, each student is required to prepare and submit his/her own solutions, including program/computer output, individually and separately. Duplication of homework solutions and codes prepared in whole or in part by someone else is not acceptable and is considered plagiarism. Collaboration of any sort on exam is prohibited and will result in a zero for the exam. Any suspicion of cheating will be reported and investigated by USC. Please see the "Academic Integrity and Conduct" section below for further details.

COURSE CALENDAR

This calendar is tentative and its details are subject to change depending on the pace of the class.

Week	Topics	Readings/ Assignments
Part I: Bandits Algorithms (formulation, learning algorithms, performance Analysis)		
1 (1/17)	Introduction to Bandit Problems	R1: Sec 1,2,4
2 (1/24)	Bandit Learning	R1: Sec 7,36
3 (1/31)	Performance Analysis of Bandit Algorithms	R1: Sec 8,36
4 (2/7)	Eluder dimension	
5 (2/14)	Adversarial Bandits	R1: Sec 11, 12 **Due: HW#1 on 2/14**
6 (2/21)	Contextual and Linear Bandits (I)	R1: Sec 18,19
7 (2/28)	Contextual and Linear Bandits (II)	R1: Sec 22, 23
8 (3/7)	Midterm Exam	
Part II: Reinforcement Learning		
9 (3/14)	Preliminaries on Markov Decision Process	R2: Ch 1.1
10 (3/21)	Spring Recess (no class)	
11 (3/28)	Planning in MDPs (Q-value iteration, Policy iteration)	R2: Ch 1.2 **Due: HW#2 on 3/28**
12 (4/4)	Model based reinforcement learning	R2: Ch 2
13 (4/11)	Strategic Explorations in RL	R2: Ch 3
14 (4/18)	Applications: Contextual Dynamic Pricing	**Due: HW#3 on 4/18**
15 (4/25)	Final Project Presentations	
16 (5/2)	Final Project Presentations	

Final Project Reports are due on May 9th, 2025 (midnight). This is the official time for our final examination.

ADDITIONAL INFORMATION

Use of AI Generators

Since creating, analytical, and critical thinking skills are part of the learning outcomes of this course, all assignments should be prepared by the student working individually or in groups. Students may not have another person or entity complete any substantive portion of the assignment. Developing strong competencies in these areas will prepare you for a competitive workplace. Therefore, using AI-generated tools is prohibited in this course, will be identified as plagiarism, and will be reported to the Office of Academic Integrity.

Open Expression and Respect for All

An important goal of the educational experience at USC Marshall is to be exposed to and discuss diverse, thought-provoking, and sometimes controversial ideas that challenge one's beliefs. In this course we will support the values articulated in the USC Marshall "Open Expression Statement" (https://www.marshall.usc.edu/open-expression-statement).

Academic Integrity

The University of Southern California is foremost a learning community committed to fostering successful scholars and researchers dedicated to the pursuit of knowledge and the transmission of ideas. Academic misconduct is in contrast to the university's mission to educate students through a broad array of first-rank academic, professional, and extracurricular programs and includes any act of dishonesty in the submission of academic work (either in draft or final form).

This course will follow the expectations for academic integrity as stated in the USC Student Handbook. All students are expected to submit assignments that are original work and prepared specifically for the course/section in this academic term. You may not submit work written by others or "recycle" work prepared for other courses without obtaining written permission from the instructor(s). Students suspected of engaging in academic misconduct will be reported to the Office of Academic Integrity.

Other violations of academic misconduct include, but are not limited to, cheating, plagiarism, fabrication (e.g., falsifying data), knowingly assisting others in acts of academic dishonesty, and any act that gains or is intended to gain an unfair academic advantage.

Academic dishonesty has a far-reaching impact and is considered a serious offense against the university. Violations will result in a grade penalty, such as a failing grade on the assignment or in the course, and disciplinary action from the university itself, such as suspension or even expulsion.

For more information about academic integrity see the student handbook or the Office of Academic Integrity's website, and university policies on Research and Scholarship Misconduct.

Please ask your instructor if you are unsure what constitutes unauthorized assistance on an exam or assignment or what information requires citation and/or attribution.

Statement on University Academic and Support Systems

Students and Disability Accommodations:

USC welcomes students with disabilities into all of the University's educational programs. The Office of Student Accessibility Services (OSAS) is responsible for the determination of appropriate accommodations for students who encounter disability-related barriers. Once a student has completed the OSAS process (registration, initial appointment, and submitted documentation) and accommodations are determined to be reasonable and appropriate, a Letter of Accommodation (LOA) will be available to generate for each course. The LOA must be given to each course instructor by the student and followed up with a discussion. This should be done as early in the semester as possible as accommodations are not retroactive. More information can be found at osas.usc.edu. You may contact OSAS at (213) 740-0776 or via email at osasfrontdesk@usc.edu.

Student Financial Aid and Satisfactory Academic Progress:
To be eligible for certain kinds of financial aid, students are required to maintain Satisfactory Academic Progress (SAP) toward their degree objectives. Visit the Financial Aid Office webpage for undergraduate- and graduate-level SAP eligibility requirements and the appeals process.

Support Systems:

Counseling and Mental Health - *(213) 740-9355 – 24/7 on call*
Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

988 Suicide and Crisis Lifeline - *988 for both calls and text messages – 24/7 on call*
The 988 Suicide and Crisis Lifeline (formerly known as the National Suicide Prevention Lifeline) provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week, across the United States. The Lifeline consists of a national network of over 200 local crisis centers, combining custom local care and resources with national standards and best practices. The new, shorter phone number makes it easier for people to remember and access mental health crisis services (though the previous 1 (800) 273-8255 number will continue to function indefinitely) and represents a continued commitment to those in crisis.

Relationship and Sexual Violence Prevention Services (RSVP) - *(213) 740-9355(WELL) – 24/7 on call*
Free and confidential therapy services, workshops, and training for situations related to gender- and power-based harm (including sexual assault, intimate partner violence, and stalking).

Office for Equity, Equal Opportunity, and Title IX (EEO-TIX) - *(213) 740-5086*
Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

Reporting Incidents of Bias or Harassment - *(213) 740-2500*
Avenue to report incidents of bias, hate crimes, and microaggressions to the Office for Equity, Equal Opportunity, and Title for appropriate investigation, supportive measures, and response.

The Office of Student Accessibility Services (OSAS) - *(213) 740-0776*
OSAS ensures equal access for students with disabilities through providing academic accommodations and auxiliary aids in accordance with federal laws and university policy.

[USC Campus Support and Intervention](#) - *(213) 740-0411*
Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

[Diversity, Equity and Inclusion](#) - *(213) 740-2101*
Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

[USC Emergency](#) - *UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call*
Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

[USC Department of Public Safety](#) - *UPC: (213) 740-6000, HSC: (323) 442-1200 – 24/7 on call*
Non-emergency assistance or information.

[Office of the Ombuds](#) - *(213) 821-9556 (UPC) / (323-442-0382 (HSC)*
A safe and confidential place to share your USC-related issues with a University Ombuds who will work with you to explore options or paths to manage your concern.

[Occupational Therapy Faculty Practice](#) - *(323) 442-2850 or* otfp@med.usc.edu
Confidential Lifestyle Redesign services for USC students to support health promoting habits and routines that enhance quality of life and academic performance.