CS282R: Topics in Machine Learning
Reinforcement Learning for Healthcare

Instructors: Finale Doshi-Velez (finale@seas.harvard.edu) and Matthieu Komorowski
TF: Omer Gottesman (gottesman@fas.harvard.edu)
Class Time and Location: MW 9-10:30 at 60 Oxford Street, Rm. 330
Office Hours: Finale: Wednesdays 3:30-4:30, MD 219, Matthieu: by appointment

Overview

Sequential decision making is at the core of many healthcare problems: a clinician observes a patient, determines a treatment, and based on the response and the patient's previous history, determines what to try next. Reinforcement learning is a formal framework for thinking about such problems. This course will first cover the fundamentals of reinforcement learning through lectures, readings, and coding assignments. Topics will include Markov decision process and partially observable Markov decision processes, planning under uncertainty, model-free and model-based reinforcement learning, function approximation in reinforcement learning, and batch reinforcement learning. Students will also engage in a semester-long project applying these techniques to clinical decision making in intensive care units.

Technical Prerequisites

Students are expected to be fluent in basic linear algebra (matrix manipulation), basic statistics (e.g. rules of expecations, importance sampling), algorithms (e.g. dynamic programming), and machine learning (at the level of CS181). From the second week, students will be working with *messy* realworld data from an intensive care unit. We will *not* be providing basic support for numpy, sklearn, etc. You will be expected to have the software engineering skills to work with data sets of 100,000+ rows. All code will be written in Python. No other languages will be supported.

Process of Course Registration and Data Access

Homework 0 Homework 0 is attached to the end of this syllabus. You should submit it on Canvas, and also email it as a pdf attachment to Omer Gottesman at gottesman@fas.harvard.edu with [CS282R: HW0] in the subject line. Homework 0 is due on midnight, Friday September 1.

Lottery The class will be limited to 20 students. A variety of factors will be taken into account in the selection process, including whether you have completed the first homework. To enter the lottery, fill out the survey below at `http://tinyurl.com/yb3z2y5n`

You will be informed by the evening of Monday, September 4 whether you have been selected. If selected, you *must* complete the following human subject certification and data access request procedure by class time on Wednesday, September 6. Otherwise, your spot will be given to someone else.

Data Access This course will involve the use of real medical data. As such you will be required to complete a human subject certification in order to participate. Once you are cleared, you will be given a cleaned up version of the data specific to this course. The process for getting access to the data is as follows:

1. Complete certification CITI "Data or Specimens Only Research" course as an MIT affiliate (not any other institution). More details are here: `https://physionet.org/works/MIMICIIClinicalDatabase/CITI_instructions.shtml`

2. Follow instructions at `https://physionet.org/works/MIMICIIClinicalDatabase/access.shtml` to sign the data use agreement and submit your CITI. Information for the form:

   - Supervisor's name: Finale Doshi-Velez
   - Supervisor's telephone number: (617) 495-2719
   - Supervisor's email address: finale@seas.harvard.edu
   - Supervisor's title: Assistant Professor

   General research area for which the data will be used: This course will be used to investigate the management of sepsis as part of Harvard CS282R, Reinforcement Learning for Healthcare, co-taught by Finale Doshi-Velez and Matthieu Komorowski.

3. Once you have completed the certification, email the pdf as an attachement to Omer Gottesman at gottesman@fas.harvard.edu with [CS282R: CITI] in the subject line.

Format, Assignments, and Assessment

The first several weeks will consist of lectures on reinforcement learning and the clinical problem. Next, we will dive into papers that are particularly relevant to reinforcement learning in batch scenarios, that is, scenarios in which a lot of data has already been collected. There will be four homework assignments (24%), readings and discussion (16%), in-class concept quizzes (10%), and a substantial semester-long project (50%).

Homework

The goal of the homework assignments is to get you familiar with basic algorithms in reinforcement learning as well as the clinical data. The work that you do will provide evaluation procedures and baselines for your semester-long project. You will have two late days to use whenever you wish in the term, except for final project write-ups. *Since the homework assignments are a basis for the final project, failure to complete the homeworks in a satisfactory manner may result in you getting kicked out of the course (staff discretion).*

What you should submit: You should submit a write-up answering the questions posed in each assignment. Your write-up should be no more than 2 pages, though you may reference plots on additional pages (please do not make your plots tiny just to make them fit). Your code should be appended to the end of the write-up. *You will be graded on the write-up only. We will not run your code.*

Collaboration: You must include the names of any people you worked with at the top of your write-up, and in what way you worked them (discussed ideas vs. team coding). If you code with others—which can be very productive!—you must have been an active participant. We may occasionally check in with groups to ascertain that everyone in the group was participating in a team-coding exercise. Your write-up must be your own.

Paper and Data Discussions

We will rotate through students presenting and discussing papers (the number of times each student presents will depend on the number of students in the class).

Presenters are responsible for reading the paper in advance and preparing a presentation (a) explaining the key ideas of the paper and (b) points for discussion. Expect to spend about half the class presenting and half the class in discussion.

Participants are responsible for reading the paper in advance and posting to Canvas with either some part of the paper you found interesting/insightful that you think is worth sharing with the class, or a question that you would like to have clarified.

Concept Checks

Throughout the semester, there will be a series of concept checks—in class quizzes to be completed on your own. These concept checks will not be announced in advance and will cover material that should be relatively straightforward (easy to work out with pen and paper). We will also monitor overall class performance on the concept checks to assess our teaching.

Semester Projects

The ideal semester project will be submitted for clinical or machine learning publication. That said, the projects will be evaluated on the quality of your research process. It is entirely okay to try out a creative idea and find it doesn't quite pan out—as long as you can explain *why*. The objective of the semester project is determine ways for to improve the management of sepsis in the ICU. You will be assigned into teams of 2-3 for this project.

Assessment will include three checkpoints (7% each): In checkpoint one, you will submit a 2-3 paragraph summary of your intended project direction, what you will achieve by the next checkpoint, and relevant references. In checkpoint two, you will submit a 2 page update which includes (a) a formal problem specification (b) preliminary results. In checkpoint three, you will submit a 2 page update which refines the problem specification and includes additional results.

The final report (30%) will be an expansion of this basic format. It is absolutely critical that your writing is clear, and that you explain *why* your ideas succeeded or failed. A series of indecipherable equations followed by dazzling plots alone will not result in high score, no matter how dazzling the plots!

Calendar

The following is a calendar of readings and assignments. Assignments will be due on Canvas.

Date	Theme	Assignments
Wednesday, August 30	Intro to RL (Sutton and Barto)	HW0 released, due Sept. 1
Wednesday, September 6	Temporal Difference (Sutton and Barto)	HW1 released
Monday, September 11	Model-Based RL (Near-Optimal Reinforcement Learning in Polynomial Time; R-MAX - A General Polynomial Time Algorithm for Near-Optimal Reinforcement Learning; A Bayesian Framework for Reinforcement Learning; A Bayesian Sampling Approach to Exploration in Reinforcement Learning	
Wednesday, September 13	Monte Carlo Approaches (Bandit based Monte-Carlo Planning; Reinforcement Learning and Simulation-Based Search in Computer Go; Doubly Robust Off-policy Value Evaluation for Reinforcement Learning; Eligibility Traces for Off-Policy Policy Evaluation)	HW2 released, HW1 due
Monday, September 18	Matthieu Komorowski: Intro to the Data, I	
Wednesday, September 20	Matthieu Komorowski: Intro to the Data, II (Continuous State-Space Models for Optimal Sepsis Treatment - a Deep Reinforcement Learning Approach; Raghu, Komorowski, Celi, Szolovits, Ghassemi)	HW3 out, HW2 due
Monday, September 25	Batch RL: Off-policy evaluation (Data-Efficient Off-Policy Policy Evaluation for Reinforcement Learning; Thomas, Brunskill) (Finale)	
Wednesday, September 27	Batch RL: Fitted-Q Iteration (Tree-Based Batch Mode Reinforcement Learning; Ernst, Geurts, Wehenkel) (Finale)	HW3 due
Monday, October 2	Batch RL: LSPI (Least-Squares Policy Iteration; Lagoudakis, Parr) (Student)	

Wednesday, October 4	Batch RL Artificial Trajectories (Batch Mode Reinforcement Learning based on the Synthesis of Artificial Trajectories; Fonteneau, Murphy, Wehenkel, Ernst) (Student)	
Wednesday, October 11	Checkpoint 1 Discussion	Checkpoint 1 due
Monday, October 16	Batch RL: DQN I (Playing Atari with Deep Reinforcement Learning; Mnih, Kavukcuoglu, Silver, Graves, Antonoglou, Wierstra, Riedmiller) (Student)	
Wednesday, October 18	Batch RL: DQN II (No reading, Taylor Killian presenting)	
Monday, October 23	Batch RL: Model-based (PILCO: A Model-Based and Data-Efficient Approach to Policy Search; Deisenroth, Rasmussen) (Student)	
Wednesday, October 25	Checkpoint 2 Discussion	Checkpoint 2 due
Monday, October 30	Policy Search () (Student)	
Wednesday, November 1	TBA - Policy Search? (Student)	
Monday, November 6	TBA - Transfer? (Student)	
Wednesday, November 8	Checkpoint 3 Discussion	Checkpoint 3 due
Monday, November 13	TBA - Policy Search or Transfer? (Student)	
Wednesday, November 15	POMDPs (Planning and acting in partially observable stochastic domains; Kaelbling, Littman, Cassandra) (Student)	
Monday, November 20	PSRs (Predictive Representations of State; Littman, Sutton, Singh) (Student)	
Monday, November 27	Final Presentations	
Wednesday, November 29	Final Presentations	
Monday, December 4		Final Write-up Due