CSE5820. Advanced Machine Learning
Reinforcement Learning

Instructor: Jinbo Bi
Tue./Thurs. 9:30pm – 10:45pm
Room: MCHU 110

The objective of this course is to introduce reinforcement learning (RL) to students including important concepts, design principles, widely used algorithms, RL with functional approximation, and generalization of RL. The course will also discuss stochastic gradient optimization algorithms that are important in understanding machine learning including RL. Another goal is for students to be able to review and understand literature on specific topics related to RL and design innovative methods to address research problems.

The course consists of lectures, programming-based and non-programming homework assignments, literature review, quizzes, and a final term project. Lectures will serve as the vehicle for the instructor to introduce concepts and knowledge. Two quizzes are used to test if basic concepts have been mastered. Homework assignments will be used for students to get profound hands-on experience by programming or experimenting with certain RL algorithms. Paper reviewing is used to improve student research capability of understanding research papers in the related fields.

Instructor: Jinbo Bi (jinbo.bi@uconn.edu)
Office: ITEB 233
Office Hours: Tuesday 11-12noon (in person or virtual https://uconn-cmr.webex.com/meet/jib10001)
Teaching Assistant: Zijie Pan (zijie.pan@uconn.edu)
Lab: ITEB 319
TA Office Hours: Tuesday 1-2pm (in person or virtual https://uconn-cmr.webex.com/meet/ajp06003)

Course Objectives:

- Learn about the core approaches and challenges of reinforcement learning
- Understand the sample complexity, generalization, approximation of the approaches
- Learn the fundamental ideas behind stochastic algorithms for large-scale optimization
- Hands-on experience with reinforcement learning algorithms to solve practical problems

Prerequisites:

CSE4820 or CSE5819
CSE3500
STAT 3025Q, or 3345Q or 3375Q, or MATH 3160

In general, students are expected to have a solid understanding of the following areas

- Algorithms: e.g., What sorting algorithms or a search tree do?
- Probability: e.g., What is Bayes rule? How do you normalize a distribution?

- Linear algebra and Calculus: e.g., What are derivatives of a multivariate function? How to compute matrix-matrix and matrix-vector products and their derivatives? What is the multivariate chain rule?
- Computer vision: Convolutional networks, object detection architectures, RNN/LSTMs
- Deep Learning: familiarity with TensorFlow or Pytorch
- Programming: e.g., What are classes and inheritance? How do you structure read data from files and how do you plot figures to visualize results using python?
- Numerical programming: e.g., How would you perform an elementwise product instead of an inner product? How do you invert a matrix?

Optional Textbooks:

- **Reinforcement Learning: An Introduction** (second edition) by Richard S. Sutton and Andrew G. Barto, ISBN-10: 0262039249

This textbook has a free online version which can be found here. It may help understand the course materials and expand the content discussed in lectures. Lectures may come with slide files, and if possible, link to textbook chapters and recent papers for students to study after lectures.

Tentative Topics and Schedule:

Recap of introductory ML and introduction of RL problems ---- 1 week
Multi-armed bandits ---- 1 week
Markov decision processes, RL problems, value iteration, policy iteration ---- 1 week
Imitation learning ---- 1 week
Dynamic programming ---- 1 week
Monte Carlo methods ---- 1 week
Temporal difference learning methods ---- 1 week
n-step bootstrapping and TD(λ) (optional) ---- 1 week
Function approximation in prediction and control, on-policy/off-policy methods ---- 2 week
Policy gradient methods (reinforce and PPO) ---- 1 week
Soft actor-critic ---- 1 week
Stochastic gradient descent, stochastic variance reduction gradient ---- 1 week
Student paper/project presentations ---- 1 week (at the first 1/3 and second 1/3 of the semester)
The above schedule is tentative, and it may adapt to a specific class scheduling. Often times, the term project is a car race or competition.

Grading:

1. Participation in-class discussion: 4% (if more than half of the in-class small quizzes are answered, 4% will be given; or otherwise 0%)
2. Homework assignments (3-4): 30%
3. In-class exam (2): 20%
4. Paper review and presentation (2): 16% (turning in review critiques/slides 4% for each presentation; actual presentation 4% for each review)
5. Final term project (1): 30%

Course Policy:

Attendance/participation includes active involvement in class discussions and presentation evaluations. Instructor may call from the class roster some students to answer questions.
For in-person sessions, computers are allowed in classroom for taking notes or any activity related to the current class meeting.
For online sessions, communications with the instructor and TA may happen in a specific timeslot or by appointments.
Students are responsible for knowing all announcements and supplements given within each class meeting and HuskyCT announcements.

Homework assignments are usually given on Monday, and will be due on the next Tuesday before class meeting time. You can have a total of 5 free late calendar days to use for homework assignments. Once late days are exhausted, no late assignments will be accepted for any reason. You are encouraged to discuss homework problems with your classmates but you must independently write your own solutions by yourself.

Two in-class exams are expected to be close-book/close-notes exams. Exam date will be informed in class at least one week ahead of time. No makeup exams will be given except for medical emergencies with an approval from the Dean of Students.

Paper review requires students to form teams (each team should have 5-6 students.) Papers will be assigned by the instructor to each team. Multiple teams may be assigned the same paper if the paper is important. Exam problems may come from some of the important papers. Each team will give a presentation on the paper they reviewed and presentation should provide summary of the paper, technical detail, experimental results if any in the paper, advantages and weaknesses of the proposed technique. However, every student needs to turn in his/her own critique in HuskyCT. The team should submit their set of slides in HuskyCT before in-class presentation. Students may expect to have two paper review sessions.

The final project will be in lieu of a final exam. The project will be done as teamwork. The project theme of this semester is focused on autonomous driving and transportation. An AWS DeepRacer competition has been set up. Students groups will train their own RL models provided from AWS using all of the virtual tracks provided by AWS. Students can equip their cars with camera, stereo camera, and/or LiDAR sensor, and use any of the state of the art RL algorithms provided by AWS and adjust the hyperparameters. The teams will be ranked according to the best lap time as shown in leaderboard. As long as the model can run, the team will receive a base score of 50 points. If a team cannot make their model run, they may only receive 20 default points. The ranking of the competed teams will receive additional scores in the following scheme: top1: +20; top 2: +15; top 3 +10; and then +5 for all other teams. After the competition, each team will present their strategies and the presentation is worth 15 points. Every student needs to write their own 2-page report summarizing what they have learned from running the RL models, which is worth 15 points.

HuskyCT:

A HuskyCT site has been set up for the class. You can access it by logging in with your NetID and password. You must use HuskyCT for submitting assignments. The instructor uses the

HuskyCT announcement to announce class materials, grades, problem clarifications, changes in class schedule, and other class announcements.

Student Responsibility and Resources:

As a member of the University of Connecticut student community, you are held to certain standards and academic policies. In addition, there are numerous resources available to help you succeed in your academic work. Review these important standards, policies and resources, which include:

- The Student Code
- Academic Integrity
- Resources on Avoiding Cheating and Plagiarism
- Copyrighted Materials
- Netiquette and Communication
- Adding or Dropping a Course
- Academic Calendar
- Policy Against Discrimination, Harassment and Inappropriate Romantic Relationships
- Sexual Assault Reporting Policy

Students with Accommodations:

Any student with a disability who needs a classroom accommodation, access to technology or other assistance in this course should contact the Center for Students with Disabilities and inform the instructor, so that arrangements can be made to accommodate the student as well as possible.

If a student is tested positive for covid or has exposure with positive cases, (s)he should follow the University policies for quarantine and is expected to self-study course materials from HuskyCT (and borrow notes from classmates or discuss with TA if necessary). Course works are still expected to be submitted on time unless a PCR test positive result is given showing the vicinity to the due date.

Academic Integrity:

You are expected to adhere to the highest standards of academic honesty. Use of published materials is allowed, but the sources should be explicitly stated in your solutions. Violations will be reviewed and sanctioned according to the University Policy on Academic Integrity. Lack of knowledge of the Student Code is not a reasonable explanation for a violation.

“Academic integrity is the pursuit of scholarly activity free from fraud and deception and is an educational objective of this institution. Academic dishonesty includes, but is not limited to, cheating, plagiarizing, fabricating of information or citations, facilitating acts of academic dishonesty by others, having unauthorized possession of examinations, submitting work for another person or work previously used without informing the instructor, or tampering with the academic work of other students.”