Machine Learning IV: Reinforcement Learning

Last updated: February 27, 2024

Instructor
Adam P Tashman
Email: apt4c@virginia.edu

Subject Area and Catalog Number: Data Science DS 7540

Year and Term: Spring 2024

Class Title. Machine Learning IV: Reinforcement Learning

Level: Graduate

Credit Type: Graded

Class Description
Reinforcement Learning is a dynamic, active area in machine learning. It is fundamentally different from supervised and unsupervised learning, as it allows an agent to learn by interacting with its environment. This allows for learning when the ground truth is unavailable, which is often the case. Even when ground truth is available, it often becomes outdated in a dynamic environment (think predicting Netflix usage before and during a pandemic). This course will introduce topics including the elements of RL, k-armed bandits, Markov Decision Processes, value functions, Policy Gradients, Q-Learning, and deep Q-Learning. The focus will be on computation and not theorem-proof (e.g., convergence guarantees).

Required Text
Reinforcement Learning 2nd edition. RS Sutton & AG Barto

Mastering Reinforcement Learning with Python. E Bilgin

Topic List (this list is subject to change)

- Foundations of Reinforcement Learning
- k-armed Bandit Problems
- Markov Decision Processes (MDPs)
- Approaches to solving MDPs:
    - > Dynamic Programming
    - > Monte Carlo
    - > Temporal-Difference Learning
- Q-Learning
- Deep Q-Learning (DQNs)
- Policy Gradients
- Safety in ML
- Reinforcement Learning with Human Feedback

Learning Outcomes

- Explain and implement methods that trade off exploration for exploitation
- Understand, compute, and improve value functions and policy gradients to optimize a policy
- Apply and contrast different methods for estimating value functions: dynamic programming, Monte Carlo simulation, temporal difference methods
- Explain how Q-Learning works and how it learns off policy
- Compare the value-based approach to the Policy Gradient approach
- Apply deep reinforcement learning techniques
- Explain methods for safely implementing reinforcement learning
- Explain methods for RL with human feedback

Delivery Mode Expectations
Weekly in-person meetings and office hours on Zoom

Required Technical Resources and Technical Components

VPN app: Cisco AnyConnect

Class Requirements

Prior to taking this course, you should meet the following course prerequisites:

- Python coding
- Machine Learning
- At least one course in Probability

Evaluation Standards and Assessments

Quizzes	Quizzes will assess student knowledge and application of topics covered in reading assignments, lecture notes, and labs.
Attendance	Student attendance is required.
Journaling	Students track their progress and learning in a Journal that is submitted to the instructor.
Labs	Labs will be implemented in Jupyter Notebooks and provide hands-on experience with concepts, RL environments, and coding.
Final Project	The final project is a large component of the course and it includes data collection, coding, algorithm use, visualization, and presentation.

Your final letter grade will be determined by the following scale:

A+	100	98.0
A	97.999	93.0
A-	92.999	90.0
B+	89.999	87.0
B	86.999	83.0
B-	82.999	80.0
C+	79.999	77.0
C	76.999	73.0

C-	72.999	70.0
D+	69.999	67.0
D	66.999	63.0
D-	62.999	60.0
F	59.999	0

Class Schedule

Mon / Wed / Fri 8:00am – 8:50am Eastern Time
Classroom: Dell 1 104

Communication & Student Response Time

Teams: We will use a Teams channel for the course. This will be a place where students can reach out to peers and instructors to ask questions and provide comments related to content and technology.

Canvas: Class announcements will be made through Canvas. High-level notes (e.g., due dates) will be posted on the Canvas home page.

Email: For questions about assignment grading or assignment submission, please email the TA.

Response Time: The instructor and TA aim to provide a response within 24 hours.

Issues that may Arise: Throughout our time together, the sooner you inform the instructor of any problem (personal or academic) that may affect your attendance or performance, the better the chance we have of solving it together.

Assignments

Quizzes (20% of grade)

All quizzes are multiple choice, with full points awarded for a correct answer, and no points awarded for an incorrect answer. Partial credit is given where appropriate. **Quizzes are closed book.**

Attendance (5% of grade)

Student attendance is required. If a student cannot attend a class, the instructor needs to be notified in advance.

Journaling (15% of grade)

The purpose of journaling is to track student learning and growth throughout this course.
The journal will be submitted over the course of the term, appending new entries.

Labs (30% of grade)

Labs will include exercises in algorithms, simulation, and machine learning. The outline of the exercise will be sketched out by the instructor to some degree. Students will fill in the code to run experiments and answer questions. **Students may collaborate with classmates but each student's code must be their own.**

Final Project (30% of grade)

The final project will have the student solve a problem using reinforcement learning, or dive deeper into an area of RL. Students can work individually or in small teams of 2 or 3. The project will consist of three components, each worth one-third of the final project grade:

1. Code
2. Presentation
3. Paper

Electronic Submission of Assignments

All assignments must be submitted electronically through Canvas by the specified due dates and times. It is crucial to complete all assigned work—failure to do so will likely result in failing the class. For late assignments, 10% of the total grade will be deducted per day, where the day means 11:59 p.m. Eastern time cutoff. After five days late, it will be marked as 0 points.

Technical Support

Technical Specifications: Computer Hardware

Operating system: Microsoft Windows 8.1 (64-bit) or Mac OS X 10.10 Minimum hard drive free space: 100 GB, SSD recommended Minimum processor speed: Intel 4th Gen Core i5 or faster
Minimum RAM: 4 GB

For the computing work, options include your machine, Rivanna, and Google Colab.

Technical Support Contacts

For issues with computing on Rivanna, please email UVA Research Computing at:
hpc-support@virginia.edu

UVA Policies

SDS Grading Policies

The standing of a graduate student in each course is indicated by one of the following grades:
A+, A, A-; B+,B, B-; C+, C, C-; D+, D, D-; F.
B- is the lowest satisfactory grade for graduate credit.

Attendance

Students are expected to attend all class sessions. Instructors establish attendance and

participation requirements for each of their courses. Class requirements, regardless of delivery mode, are not waived due to a student's absence from class. Instructors will require students to make up any missed coursework and may deny credit to any student whose absences are excessive. Instructors must keep an attendance record for each student enrolled in the course to document attendance and participation in the class.

University Email Policies

Students are expected to check their official UVA email addresses on a frequent and consistent basis to remain informed of University communications, as certain communications may be time sensitive. Students who fail to check their email on a regular basis are responsible for any resulting consequences.

End-of-Class Evaluations

Students are expected to complete the online end-of-class evaluation. As the semester comes to a close, students will receive an email with instructions for completing this. Student feedback will be very valuable to the school, the instructor, and future students. We ask that all students please complete these evaluations in a timely manner. Please be assured that the information you submit online will be anonymous and kept confidential.

University of Virginia Honor System

All work should be pledged in the spirit of the Honor System at the University of Virginia. The instructor will indicate which assignments and activities are to be done individually and which permit collaboration. The following pledge should be written out at the end of all quizzes, examinations, individual assignments and papers: "I pledge that I have neither given nor received help on this examination (quiz, assignment, etc.)." The pledge must be signed by the student. For more information, visit www.virginia.edu/honor.

Special Needs

It is my goal to create a learning experience that is as accessible as possible. If you anticipate any issues related to the format, materials, or requirements of this course, please meet with me outside of class so we can explore potential options. Students with disabilities may also wish to work with the StudePnt Disability Access Center to discuss a range of options to removing barriers in this course, including official accommodations. Please visit their website for information on this process and to apply for services online: sdac.studenthealth.virginia.edu. If you have already been approved for accommodations through SDAC, please send me your accommodation letter and meet with me so we can develop an implementation plan together.