Habib University
shaping futures

Introduction to Reinforcement Learning

CS 352 L1
Spring Semester 2023

Course Information

Class Location: FF-N220

Class Meeting Time(s): Wed Fri (08:30 AM-09:45 AM)

Course Prerequisites: CS/CE 102/171,MATH 310
Course Prerequisites: CS/CE 102/171,MATH 310CS 102, Probability and Statistics

Hardware/Software Prerequisites (if any): LaTeX, Python 3
Content Area:

Instructor Information

Instructor: Mohammad Shahid Shaikh
Office Location: C-212
Email: shahid.shaikh@sse.habib.edu.pk
Office Hours: To be decided.

Course Description

This is an introductory but mathematically rigorous course on Reinforcement Learning (RL). RL is a machine learning (RL) paradigm that is distinct from supervised and unsupervised learning. In RL, the learning agent modifies its action within a learning environment so as to maximize the accumulated

reward. In this sense, among the three learning paradigms–supervised, unsupervised and RL–RL most closely mimics the way human beings and animals learn to perform various tasks.

Course Aims

In this course, students will be introduced to the modern RL framework and how to frame a given ML problem within this framework. They will also be introduced to variations of the basic framework and will learn commonly used techniques for solving RL problems using custom-made and ready-made software packages.

Course Learning Outcomes (CLOs)

By the end of this course, the students will be able to:

1. Define reinforcement learning and distinguish it from other machine learning paradigms.
2. Develop models of some well-known machine learning problems using the RL framework.
3. Describe and analyze some well-known algorithms for solving problems given in the standard RL framework.
4. Implement common RL algorithms using Python programming language and off-the-shelf solutions.

Mode of Instruction

In-person, face-to-face.

Required Texts and Materials

Reinforcement Learning: An Introduction
ISBN: 978-0262193986
Edition: second
Authors: Richard S. Sutton and Andrew G. Barto
Publisher: MIT Press, Cambridge
Publication Date: 2018
Available for free download at:
https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf

Reading material provide by the instructor.

Assessments

Assessment	Percentage	Description
Assignments	30%	Three assignments, worth 10% each.
Midterm	20%	One midterm exam in the 8th week.
Quizzes	25%	Six quizzes worth 5% each. The best five will count.
Project	25%	Analysis, modeling, solution, and software implementation of an RL application.

Grading Scale

Letter Grade	GPA Points	Percentage
A+	4.00	[95-100]
A	4.00	[90-95)
A-	3.67	[85-90)
B+	3.33	[80-85)
B	3.00	[75-80)
B-	2.67	[70-75)
C+	2.33	[67-70)
C	2.00	[63-67)
C-	1.67	[60-63)
F	0.00	[0, 60]

Note: [a, b) is a range of numbers from a to b where a is included in the range and b is not.

Late Submission Policy

No late submissions allowed.

Week-Wise Schedule (Tentative)

Week	Topics(s)	Reading(s)	Remarks
1	Elements of Reinforcement Learning	Chapter 1	

2	Tabular methods for elementary problems	Chapter 2	Quiz #1
3	The standard Reinforcement Learning framework	Chapters 2 & 3	Assignment #1 released
4	Review of Markov decision processes	Chapter 3	Quiz #2
5	Policy evaluation and improvement	Chapter 4	Assignment #1 due
6	Policy iteration	Chapter 4	Quiz #3
7	Dynamic Programming and Value iteration	Chapter 4	Assignment #2 released
8	Approximate Dynamic Programming (Monte Carlo Methods)	Chapter 5	Midterm Exam
9	Approximate Dynamic Programming (continued)	Chapter 5	Assignment #2 due
10	Temporal-Difference (TD) Method	Chapter 6	Quiz #4
11	TD method (continued)	Chapter 6	Assignment #3 released
12	State-Action-Reward- State-Action (SARSA) method	Chapter 7	Quiz #5
13	Q-learning	Chapter 7	Assignment #3 due
14	Tabular Methods and Function Approximation	Reading material provided by the instructor	Quiz #6
15	Project Viva and presentations		Project viva and presentations

Attendance Policy

As per the University's requirement.

Academic Integrity (Updated)

Each student in this course is expected to abide by the Habib University Student Honor Code of Academic Integrity. Any work submitted by a student in this course for academic credit will be the

student's own work.

Scholastic dishonesty shall be considered a serious violation of these rules and regulations and is subject to strict disciplinary action as prescribed by Habib University regulations and policies. Scholastic dishonesty includes, but is not limited to, cheating on exams, plagiarism on assignments, and collusion.

a. Plagiarism: Plagiarism is the act of taking the work created by another person or entity and presenting it as one's own for the purpose of personal gain or of obtaining academic credit. As per University policy, plagiarism includes the submission of or incorporation of the work of others without acknowledging its provenance or giving due credit according to established academic practices. This includes the submission of material that has been appropriated,// bought, received as a gift, downloaded, or obtained by any other means. Students must not, unless they have been granted permission from all faculty members concerned, submit the same assignment or project for academic credit for different courses.
b. Cheating: The term cheating shall refer to the use of or obtaining of unauthorized information in order to obtain personal benefit or academic credit.
c. Collusion: Collusion is the act of providing unauthorized assistance to one or more person or of not taking the appropriate precautions against doing so.

All violations of academic integrity will also be immediately reported to the Student Conduct Office.

You are encouraged to study together and to discuss information and concepts covered in lecture and the sections with other students. You can give "consulting" help to or receive "consulting" help from such students. However, this permissible cooperation should never involve one student having possession of a copy of all or part of work done by someone else, in the form of an e-mail, an e-mail attachment file, a diskette, or a hard copy.

Should copying occur, the student who copied work from another student and the student who gave material to be copied will both be in violation of the Student Code of Conduct.

If you wish to use generative-AI tools to complete any of your assessments, you must first obtain permission from your course instructor. AI generated work will not be accepted in all classes or even all assessments. The instructor's permission is required. If the permission is granted, you should declare its use and properly cite the source of the generated content. Failing to identify AI written or assisted work is academic dishonesty and will be treated as any case of plagiarism by the university.

The principle for academic integrity is that your submissions must be substantially your own work and that any work that is not originally your thought must be identified and credited. If the use of AI tools is prohibited in the course, respect the rules and do not use these tools for assessments. The fundamental purpose of assessment is to learn, synthesize information and explain new connections and interpretations that arise from your secondary research. Be aware that unauthorized use of AI

tools for assessments can result in a conduct case being filed. This can have serious consequences for your academic standing and future career opportunities.

During examinations, you must do your own work. Talking or discussion is not permitted during the examinations, nor may you compare papers, copy from others, or collaborate in any way. Any collaborative behavior during the examinations will result in failure of the exam, and may lead to failure of the course and University disciplinary action.

Penalty for violation of this Code can also be extended to include failure of the course and University disciplinary action.

Program Learning Outcomes (For Administrative Review)

Upon graduation, students will have the following abilities:

- PLO 1: Analysis: Analyse a given situation and reduce it to one or more problems that can be solved via computer intervention.
- PLO 2: Design: Design one or more computer-based solutions of a given problem and select the solution that is best under the circumstances.
- PLO 3: Programming: Program a given solution in a variety of programming languages belonging to different paradigm.
- PLO 4: Implementation: Design and implement software systems of varying complexity.
- PLO 5: Tools: Work with the latest tools that support development, e.g., IDE's, version control systems, debuggers, profilers, and continuous build systems.
- PLO 6: Self-learning: Research, learn, and apply requirements needed to implement a solution for a given high level problem description.
- PLO 7: Ethics and Awareness: Foresee both impact and possible ramifications of computing practices.
- PLO 8: Communication and Teamwork: Work effectively in inter-disciplinary teams.

Program Learning Outcomes (PLOs) mapped to Course Learning Outcomes (CLOs)	
	CLOs of the course are designed to cater following PLOs:

	PLO 1: PLO 2:				
	Distribution of CLO weighs for each PLO				
	CLO 1	CLO 2	CLO 3	CLO 4	CLO 5
PLO 1		X	X	X	
PLO 2			X	X	
PLO 6			X	X	

Mapping of Assessments to CLOs

Assignments	CLO #01	CLO #02	CLO #03	CLO #04	CLO #05

Recording Policy

Only asynchronous and synchronous online sessions will be recorded and uploaded on our Video Management System (Panopto). Link to the folder of recordings will be available to all students. Hyflex classes might be recorded if faculty deems it appropriate.

Accommodations for Students with Disabilities

In compliance with the Habib University policy and equal access laws, I am available to discuss appropriate academic accommodations that may be required for student with disabilities. Requests for academic accommodations are to be made during the first two weeks of the semester, except for unusual circumstances, so arrangements can be made. Students are encouraged to register with the Office of Academic Performance to verify their eligibility for appropriate accommodations.

Inclusivity Statement

We understand that our members represent a rich variety of backgrounds and perspectives. Habib University is committed to providing an atmosphere for learning that respects diversity. While working together to build this community we ask all members to:

- share their unique experiences, values and beliefs
- be open to the views of others
- honor the uniqueness of their colleagues
- appreciate the opportunity that we have to learn from each other in this community
- value each other's opinions and communicate in a respectful manner
- keep confidential discussions that the community has of a personal (or professional) nature
- use this opportunity together to discuss ways in which we can create an inclusive environment in this course and across the Habib community

Office Hours Policy

Every student enrolled in this course must meet individually with the course instructor during course office hours at least once during the semester. The first meeting should happen within the first five weeks of the semester but must occur before midterms. Any student who does not meet with the instructor may face a grade reduction or other penalties at the discretion of the instructor and will have an academic hold placed by the Registrar's Office.