

HABIB UNIVERSITY

CS-4xx: Introduction to Reinforcement Learning Spring 2023

Instructor:	Dr. M. Shahid Shaikh
Contact:	shahid.shaikh@sse.habib.edu.pk , Ext. 5222
Class Location:	N-220
Class Meeting Time(s):	Wednesdays and Fridays 8:30 am to 9:45 am
Office Hours:	Wednesdays and Fridays 10:00 am to 11:00 am
Course Prerequisites:	Math 310, Probability and Statistics and CS 102, Data Structures and Algorithm
Software Prerequisites:	LaTeX, Python 3.
Hardware Prerequisites:	Computer with mic, camera, and Internet connection. (for online classes)

I. Rationale:

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.

II. Course Aims and Outcomes:

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.

CLOs to ECE PLOs mapping

	CLOs of the course are designed to cater to PLO 1: Engineering Knowledge, PLO 2: Problem Analysis PLO 3: Design/Development of Solutions PLO 5: Modern Tool Usage			
	Distribution of CLO weights for each PLO			
PLOs	CLO 1	CLO 2	CLO 3	CLO 4
PLO 1				
PLO 2				
PLO 3				
PLO 5				

CLOs to CS PLOs mapping

	CLOs of the course are designed to cater to PLO 1: Analysis PLO 2: Design PLO 6: Self Learning			
	Distribution of CLO weights for each PLO			
PLOs	CLO 1	CLO 2	CLO 3	CLO 4
PLO 1				
PLO 2				
PLO 6				

I. Format and Procedures:

The course will be delivered through lectures. Assessment will consist of assignments, quizzes, exams, and a project. Assignments will consist of questions based on the theory learned in the class and software implementation of the theoretical ideas and use of relevant software tools.

IV. Course Requirements:

1. Class participation policy: Students are highly encouraged to ask questions and participate in the discussion on the relevant topics.

2. Course readings:

- i. Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction*, Second Edition, MIT Press, Cambridge, MA, 2018.

Available for free download at:

<https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPLBook2ndEd.pdf>

- ii. Reading material provide by the instructor.

V. Grading Procedures:

Assessment	Percentage	Description
Assignments	30%	Three assignments, worth 10% each.
Midterm	20%	One midterm exam in the 8 th 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.

VI. Mapping of Assessments to CLOs

	CLO #01	CLO #02	CLO #03	CLO #04
Assignment 1				
Assignment 2				
Assignment 3				
Quiz No 01				
Quiz No 02				
Quiz No 03				
Quiz No 04				
Quiz No 05				
Quiz No 06				
Mid-Term Exam				
Course Project				

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)

VI. Attendance Policy:

Habib University requires that all students must maintain at least 85% for each class in which they are registered. Non-compliance with minimum attendance requirements will result in automatic failure of the course and may require the student to repeat the course when next offered. This policy is at a minimum. Departments, schools, and individual faculty members may alter this policy to include stronger attendance requirements and/or implement them for all levels of students. It is the responsibility of the student to keep track of their own attendance and speak with their faculty member or the Office of the Registrar for any clarification.

VII. 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 a 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.

VIII. 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

IX. Office hours:

Office hours are your chance to meet the instructor for closer personalized interaction. This may be to discuss specific queries on particular course work, for a general discussion on the course, your overall performance in the course, or any other course-related aspects. Instructors are also happy to discuss more general academic matters, e.g. your eventual career goals, the state of the art in their research area, ideas for research projects, and so on. While instructors welcome interaction with the students, office hours are time that they dedicate in their schedule for the above interaction.

X. Academic Integrity

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.

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.

CHEATING: The term cheating shall refer to the use of or obtaining of unauthorized information in order to obtain personal benefit or academic credit.

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 the lectures.

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

XI. 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