SJSU SAN JOSÉ STATE UNIVERSITY

College of Science · Computer Science

Reinforcement Learning and Sequential Decision Making

CS 272

Spring 2026 Section 01 In Person 3 Unit(s) 01/22/2026 to 05/11/2026 Modified 01/19/2026

Course Description and Requisites

Introduction to reinforcement learning, deep reinforcement learning, other online learning algorithms, and their applications.

Prerequisite(s): CS 157A. Allowed Declared Major: Computer Science MS, Bioinformatics MS, and Data Science MS.

Letter Graded

Classroom Protocols

- This course is conducted in person, which requires physical attendance for a significant portion of its delivery.
- Students are requested to use the Canvas message function to contact the instructor. Private messages sent to the instructor's email address get lost due to the large volume of emails received.
  - The instructor does not write messages after normal business hours, on weekends, or on holidays.
  - Reviewing code for the homework and technical troubleshooting should be done during office hours. Never send your entire code for an assignment to the instructor. The instructor will not fix the bugs in your code.

Program Information

Diversity Statement - At SJSU, it is important to create a safe learning environment where we can explore, learn, and grow together. We strive to build a diverse, equitable, inclusive culture that values, encourages, and supports students from all backgrounds and experiences.

Course Learning Outcomes (CLOs)

Upon successful completion of this course, students will be able to:

- Distinguish different types of reinforcement learning algorithms and when to use them.
- Describe the benefits and potential challenges of deep reinforcement learning.
- Apply reinforcement learning algorithms to real-world problems.
- Analyze and evaluate the performance of reinforcement algorithms.
- Create a reinforcement learning project to solve a social or technical issue.

Course Materials

Textbook:

- Richard S. Sutton and Andrew G. Barto, Reinforcement learning: An introduction (Second edition), MIT press, 2018. [Available for free on the authors' website]
- Alexander Zai and Brandon Brown, Deep Reinforcement Learning in Action, (https://learning.oreilly.com/home/) ISBN 9781617295430, 2020. [Available for free through your SJSU account]

Supplemental Textbooks:

- Michael A. Nielsen, Neural Networks and Deep Learning, Determination Press, 2015.
- Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong, Mathematics for Machine Learning, Cambridge University Press, 2020.

Technology

- Python development environment with the following packages (It is highly recommended to use an isolated Python workspace, such as Anaconda env.)
  - Gymnasium (https://gymnasium.farama.org/index.html (https://gymnasium.farama.org/index.html))
  - PettingZoo (https://pettingzoo.farama.org/ (https://pettingzoo.farama.org/))
  - stable-baselines3 (https://stable-baselines3.readthedocs.io/en/master/ (https://stable-baselines3.readthedocs.io/en/master/))
  - Ray RLlib (https://docs.ray.io/en/latest/rllib/index.html (https://docs.ray.io/en/latest/rllib/index.html))
- All demos assume a UNIX-like OS environment.

Course Requirements and Assignments

Item	Percentage
Wednesday Quiz 1 - 8 (wq)	15% (The worst three will be dropped. 3% Each)

Item	Percentage
Programming Assignments 1 - 3 (pa)	30% (10% Each)
Exam 1	20%
Exam 2	20%
Final Project	15%

Wednesday Quizzes

There will be a small quiz at the beginning of the class (almost) every Wednesday. (Check the Syllabus on Canvas for the schedule.)

The purpose of the Wednesday quizzes is to confirm your understanding of the fundamental concepts in RL right after each class. Most quizzes are multiple-choice or True-False questions and do not involve complicated calculations.

No retake or makeup opportunity will be given since the solutions will be discussed in class. The worst three (3) quizzes will be dropped at the end of the semester. If you need special accommodations (due to health concerns or other unforeseen circumstances), please speak with the instructor.

Programming Assignments (PA)

Submissions may be graded based on automated software testing. The specifications will be given in class and in the form of code templates.

Exams

Two exams will be conducted during the regular class hours. A tentative schedule will be given in the course schedule below. No make-up exams except in case of verifiable emergency circumstances.

✓ Grading Information

Extra-credits and Reworks

No extra-credit assignments or rework opportunities will be given.

Late Submission

No late submissions will be accepted.

Missed Assignments or Exams

When students need to miss an assignment deadline or exam due to health conditions or any other emergency, it should be reported within ONE week after the due date.

Final Grade Table

Total Grade	Letter Grade
97% and above	A plus
93% to 96%	A
90% to 92%	A minus
87% to 89%	B plus
83% to 86%	B
80% to 82%	B minus
77% to 79%	C plus
73% to 76%	C
70% to 72%	C minus
67% to 69%	D plus
65% to 66%	D
60% to 64%	D minus
59% and below	F

University Policies

Per University Policy S16-9 (PDF) (http://www.sjsu.edu/senate/docs/S16-9.pdf), relevant university policy concerning all courses, such as student responsibilities, academic integrity, accommodations, dropping and adding, consent for recording of class, etc. and available student services (e.g. learning assistance, counseling, and other resources) are listed on the Syllabus Information (https://www.sjsu.edu/curriculum/courses/syllabus-info.php) web page. Make sure to visit this page to review and be aware of these university policies and resources.

Course Schedule

Date	Topic	Reference	Note
1/26	Overview		
1/28	What is Learning?		

Date	Topic	Reference	Note
2/2	Markov Decision Processes	S&B Chap 3	
2/4	Policies and Value Functions	S&B Chap 3	wq1: Learning Concepts
2/9	Dynamic Programming	S&B Chap 4	
2/11	Coding: DP		wq2: MDP; PA1 due
2/16	Try and Error: MAB	S&B Chap 2	
2/18	MC	S&B Chap 5	wq3: Value functions and DP
2/23	TD	S&B Chap 6	
2/25	Policy Gradient Methods	S&B Chap 13	wq4: MC
3/2	Coding: Policy Gradient Methods		
3/4	Exam Review		wq5: TD
3/9	Exam 1		PA2 due
3/11	Coding: Gymnasium		
3/16	Approximation		
3/18	Coding: Linear Approximation	S&B Chap 10	wq6: PG
3/23	Coding: Custom Environment		
3/25	Deep Learning		wq7: Approximation; Project group due
3/30	No class (Spring Recess)		
4/1	No class (Spring Recess)		
4/6	Deep RL: DQN	S&B DRL in Action Chap 3	
4/8	Deep RL: DQN + Implementation		
4/13	Deep RL: A3C, PPO		
4/15	Coding: stable-baselines3	"Empirical Design in Reinforcement Learning" arXiv:2304.01315 [cs]	wq8: DRL; PA3 due

Date	Topic	Reference	Note
4/20	Project discussion		
4/22	Integrating Learning and Planning	S&B Chap 8	
4/27	Advanced topics in DRL: LLM with DRL		
4/29	Exam Review		
5/4	Exam 2		
5/6	Project presentation		Final project due
5/11	Project presentation		