Course Description
This course introduces the theory and practice of reinforcement learning and decision-making techniques, with applications in numerous fields, such as communications, computing, control, and caching (4C) systems, Internet of Things, cyber physical social systems, green communications, quantum communications and computing, smart grid systems, public safety systems, positioning, navigation, and timing solutions, and others. This course will introduce the students to the learning techniques, where an autonomous agent takes intelligent actions and interacts with the environment in a dynamic manner. Understanding the importance and challenges of artificial intelligent agents that make decisions is of vital importance, with more and more companies interested in interactive agents and intelligent decision-making. The course will cover tabular solution methods, multi-armed bandit problems, gradient bandit algorithms, finite Markov Decisions Processes (MDP), policies and value functions, dynamic programming, policy improvement, policy iteration, Monte Carlo methods, SARSA, Q-learning, temporal difference learning, n-step bootstrapping, on-policy prediction with approximation, on-policy control with approximation, and off-policy methods with approximation, and policy gradient methods. The audience of this course can be undergraduate and graduate students from any engineering department in the School of Engineering, UNM, including the Computer Science Department, and students from the Mathematics & Statistics, and Physics and Astronomy Departments, UNM.

Course Objectives
At the completion of the course, students will be able to
1. **Know** the fundamental reinforcement learning theory and how to apply it to real-world problems.
2. **Understand** how reinforcement learning relates to machine learning, deep learning, supervised and unsupervised learning.
3. **Define** problems as Markov Decision Processes.
4. **Formulate** decision-making problems, set up and run computational experiments, and evaluate the results from the lab experiments.
5. **Explain** when and how to apply model-based versus model-free learning methods.
6. **Describe** approaches for balancing exploration and exploitation during learning to adapt the convergence time in realistic applications.
7. **Identify** the strengths, weaknesses, and tradeoffs of different reinforcement learning algorithms.
Prerequisites and Co-requisites
Good knowledge of linear algebra, calculus, probability theory, statistics, and Python programming. This course is part of the AI courses offered in the ECE department: Machine Learning, Reinforcement Learning, Advanced Machine Learning, Network Economics (Game Theory), and Deep Learning. Students interested in the AI field, are encouraged to take this series of courses.

Technical Skills
In order to participate and succeed in this class, you will need to be able to interact with the Jupyter Notebook. Instructions will be provided in the first lecture of the course and during the instructor’s office hours, if additional help is needed.

Textbook and Supplemental Materials

Required Textbooks:


Recommended and/or Optional Textbooks:


Coursework, Participation, and Grading Procedures

Homework (30%): There will be biweekly homework, including review questions, problems, lab exercises to be implemented through Jupyter Notebook. The homework will be uploaded at UNM Canvas and will be timed. You will have one or multiple shots, depending on the assignment, to submit your answer directly at UNM Canvas.

Mid Term Exam (35% for students in ECE 402): The mid term exam is open book. No additional material will be used. If a student must miss an exam, and if that student has a written, verifiable, legitimate (health-related issues) excuse for the absence, then the student must contact the instructor prior to the Exam Date and provide necessary written evidence for the absence (evidence for health-related problems may be presented later in consultation with the instructor). The instructor reserves sole discretion in this matter of providing the opportunity for a makeup Exam. This opportunity is available only once during the semester. An undergraduate student will have the option to select among the midterm exam and the research project. The undergraduate student will be informed by the instructor that the research project requires novel research to be presented and the student’s choice among the midterm and the research project will be documented.

Research Project (35% for students in ECE 502): The students will iterate over their research project proposals, setting individual goals and the evaluation criteria will be set by the instructor, over the first four modules (i.e., four weeks) of the course. The goal of the project will be to design an RL autonomous agent in a space selected by the student, e.g., communications, computing, control, and caching (4C) systems, Internet of Things,
cyber physical social systems, green communications, quantum communications and computing, smart grid systems, public safety systems, and others, formulate the learning goal, the reward function, select the most appropriate RL algorithm in terms of computational complexity, and provide comparative evaluation to other RL algorithms or non-RL approaches (e.g., game-theoretic approaches) to show the efficiency and robustness of the learning process. Once the students and the instructor have determined an appropriately scoped project and evaluation criteria, the students will proceed with the implementation of the project, which should be delivered by the end of the semester.

Final Exam (35%): The final exam is open book. No additional material will be used. If a student must miss an examination, and if that student has a written, verifiable, legitimate (health-related issues) excuse for the absence, then the student must contact the instructor prior to the Exam Date and provide necessary written evidence for the absence (evidence for health-related problems may be presented later in consultation with the instructor). The instructor reserves sole discretion in this matter of providing the opportunity for a makeup Exam. This opportunity is available only once during the semester.

Participation (0%): Each student is responsible for all of the information (including announcements and handouts) presented in class. Traditionally, poor performance in this class has been closely related to poor attendance. Your participation is voluntary. Students are encouraged to contact the instructor and the teaching assistant during the office hours at instructor’s office.

The grading (both of the homework, the midterm exam, the research project, and the final exam) will be performed by the instructor and not by the teaching assistant. The TA will accept questions during the announced office hours. The students are encouraged to contact first the instructor for questions, and if the instructor’s schedule is busy then contact the TA. Following this pattern, you help the instructor understand the difficulties that the students face during the lectures and adjust accordingly. Your active participation is highly encouraged.

Teaching Assistant: TBA

Grading Scale

<table>
<thead>
<tr>
<th>Points (%)</th>
<th>Grade</th>
<th>Points (%)</th>
<th>Grade</th>
<th>Points (%)</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>[93,98)</td>
<td>A</td>
<td>[78,80]</td>
<td>C+</td>
<td>[60,63]</td>
<td>D-</td>
</tr>
<tr>
<td>[90,93)</td>
<td>A-</td>
<td>[73,78]</td>
<td>C</td>
<td>[0,60]</td>
<td>F</td>
</tr>
<tr>
<td>[88,90)</td>
<td>B+</td>
<td>[70,73]</td>
<td>C-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Permitted Collaboration

The following items are encouraged and allowed at all times for all students in this class:

• Discussion of material covered during lecture or in handouts
• Discussion of the requirements of an assignment
• Discussion of the use of tools
• Discussion of general approaches to solving problems
• Discussion between a student and the instructor for the course

Unpermitted Collaboration

All submissions must represent original, independent work. Some examples of activities that do not represent original work include:

• Copying solutions from others. In particular, do not ask anyone to provide a copy of his or her solution or, conversely, give a solution to another student who requests it.
• Studying another student's solution. Do not read another solution submission whether in electronic or printed form, even to "check answers."

Course Schedule

Revisit the course schedule during the semester as timing changes will be made on the fly based on your performance.

<table>
<thead>
<tr>
<th>Class Modules</th>
<th>Tentative Topics</th>
</tr>
</thead>
</table>
| Module 1      | • Introduction to Reinforcement Learning  
                • Usage of Jupyter Notebook  
                • Elements of Reinforcement Learning  
                • Limitations of Reinforcement Learning  
                Read S&B Chapter 1 |
| Module 2      | • Multi-armed bandits  
                • Action-value methods  
                • Tracking a non-stationary problem  
                • Optimistic Initial Values  
                Read S&B Chapters 2.1-2.6 |
| Module 3      | • Upper-confidence-bound action selection  
                • Gradient bandit algorithms  
                • Contextual bandits  
                Read S&B Chapters 2.7-2.10 |
| Module 4      | • Finite Markov Decision Processes (MDP)  
                • Agent-environment interface  
                • Goals and rewards  
                • Returns and Episodes  
                Read S&B Chapters 3.1-3.4 |
<p>| Module 5      | • Unified notation for episodic and continuing tasks |</p>
<table>
<thead>
<tr>
<th>Module</th>
<th>Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Dynamic programming, Policy evaluation, Policy improvement, Policy iteration, Value iteration</td>
</tr>
<tr>
<td>7</td>
<td>Asynchronous dynamic programming, Generalized policy iteration, Efficiency of dynamic programming</td>
</tr>
<tr>
<td>8</td>
<td>Monte Carlo methods, Monte Carlo prediction, Monte Carlo estimation of action values, Monte Carlo control, Questions &amp; answers – Midterm exam (ECE 402)</td>
</tr>
<tr>
<td>9</td>
<td>Off-policy prediction via importance sampling, Incremental implementation, Off-policy Monte Carlo control</td>
</tr>
<tr>
<td>10</td>
<td>Temporal-Difference (TD) Learning, TD prediction, Advantages of TD prediction methods, Optimality of TD, SARSA, Q-learning</td>
</tr>
<tr>
<td>11</td>
<td>Expected SARSA, Maximization bias and double learning, Game, afterstates, and other special cases</td>
</tr>
<tr>
<td>12</td>
<td>N-step bootstrapping, Planning and learning with tabular methods</td>
</tr>
<tr>
<td>13</td>
<td>On-policy prediction with approximation, Value-function approximation, Stochastic-gradient &amp; semi-gradient methods, Linear methods, Non-linear function approximation</td>
</tr>
<tr>
<td>14</td>
<td>Least-squares TD, Memory-based function approximation, On-policy control with approximation</td>
</tr>
<tr>
<td>15</td>
<td>Off-policy methods with approximation, Policy gradient methods</td>
</tr>
<tr>
<td>16</td>
<td>Questions &amp; Answers – Final Exam (ECE 502 and optional choice for ECE 402)</td>
</tr>
</tbody>
</table>

Read S&B Chapters 3.5-3.8, 4.1-4.4, 5.1-5.4, 5.5-5.10, 6.1-6.5, 6.6-6.9, 7-8, 9.1-9.7, 9.8-9.12, 10, 11, 13.
COVID-19 Health and Awareness

UNM is a mask friendly, but not a mask required, community. To be registered or employed at UNM, Students, faculty, and staff must all meet UNM's Administrative Mandate on Required COVID-19 vaccination. If you are experiencing COVID-19 symptoms, please do not come to class. If you have a positive COVID-19 test, please stay home for five days and isolate yourself from others, per the Centers for Disease Control (CDC) guidelines. If you do need to stay home, please communicate with me at eirini@unm.edu or (505) – 277 – 5501; I can work with you to provide alternatives for course participation and completion. UNM faculty and staff know that these are challenging times. Please let us know that you need support so that we can connect you to the right resources and please be aware that UNM will publish information on websites and email about any changes to our public health status and community response.

Student Health and Counseling (SHAC) at (505) 277-3136. If you are having active respiratory symptoms (e.g., fever, cough, sore throat, etc.) AND need testing for COVID-19; OR If you recently tested positive and may need oral treatment, call SHAC.

Accommodation Statement

In accordance with University Policy 2310 and the Americans with Disabilities Act (ADA), academic accommodations may be made for any student who notifies the instructor of the need for an accommodation. It is imperative that you take the initiative to bring such needs to the instructor’s attention, as I am not legally permitted to inquire. Students who may require assistance in emergency evacuations should contact the instructor as to the most appropriate procedures to follow. Contact Accessibility Resource Center at 277-3506 or arcsrvs@unm.edu for additional information.

UNM is committed to providing courses that are inclusive and accessible for all participants. As your instructor, it is my objective to facilitate an accessible classroom setting, in which students have full access and opportunity. If you are experiencing physical or academic barriers, or concerns related to mental health, physical health and/or COVID-19, please consult with me after class, via email/phone or during office hours. You are also encouraged to contact Accessibility Resource Center at arcsrvs@unm.edu or by phone 277-3506.

Title IX Statement

In an effort to meet obligations under Title IX, UNM faculty, Teaching Assistants, and Graduate Assistants are considered “responsible employees” by the Department of Education (see pg 15 - http://www2.ed.gov/about/offices/list/ocr/docs/qa-201404-title-ix.pdf). This designation requires that any report of gender discrimination which includes sexual harassment, sexual misconduct and sexual violence made to a faculty member, TA, or GA must be reported to the Title IX Coordinator at the Office of Equal Opportunity (oeo.unm.edu). For more information on the campus policy regarding sexual misconduct, see: https://policy.unm.edu/university-policies/2000/2740.html

Academic Integrity Statement

Each student is expected to maintain the highest standards of honesty and integrity in academic and professional matters. The University reserves the right to take disciplinary action, up to and including dismissal, against any student who is found guilty of academic dishonesty or otherwise fails to meet the standards. Any student judged to
have engaged in academic dishonesty in course work may receive a reduced or failing grade for the work in question and/or for the course (See Section: Unpermitted Collaboration). Academic dishonesty includes, but is not limited to, dishonesty in quizzes, tests, or assignments; claiming credit for work not done or done by others; hindering the academic work of other students; misrepresenting academic or professional qualifications within or without the University; and nondisclosure or misrepresentation in filling out applications or other University records.

**Citizenship and/or Immigration Status**

All students are welcome in this class regardless of citizenship, residency, or immigration status. Your professor will respect your privacy if you choose to disclose your status. As for all students in the class, family emergency-related absences are normally excused with reasonable notice to the professor, as noted in the attendance guidelines above. UNM as an institution has made a core commitment to the success of all our students, including members of our undocumented community. The Administration’s welcome is found on our website: http://undocumented.unm.edu/.

**Credit-hour statement**

This is a three credit-hour course. Class meets for two 65-minute sessions of direct instruction for fifteen weeks during the Fall 2023 semester. Please plan for a minimum of eight hours of out-of-class work (or homework, study, assignment completion, and class preparation) each week.

**Respectful and Responsible Learning**

We all have shared responsibility for ensuring that learning occurs safely and equitably. UNM has important policies to preserve and protect the academic community, especially policies on student grievances (Faculty Handbook D175 and D176), academic dishonesty (FH D100), and respectful campus (FH CO9). These are in the Student Pathfinder (https://pathfinder.unm.edu) and the Faculty Handbook (https://handbook.unm.edu). Please ask for help in understanding and avoiding plagiarism or academic dishonesty, which can both have very serious consequences.