

**Information Technology**  
**Course Number: 22:544:637**  
**Course Title: Reinforcement Learning**

## **COURSE DESCRIPTION**

The course provides an introduction to the theory and applications of Reinforcement Learning (RL). RL is an effective approach within Machine Learning (ML) suited for addressing complex business problems involving Sequential Optimal Decision-Making under Uncertainty. It stands out from other ML branches, since it enables adaptive decision-making in AI systems, continuously striving for optimal results.

RL is experiencing substantial recognition through its successful applications in Data Science and Machine Learning, Robotics and Industrial Automation, Healthcare and Clinical Trials, Media Advertising, Recommender Systems, and Finance.

---

## **COURSE MATERIALS**

The course notes will be posted on the Rutgers Box server

### **Recommended Texts:**

- Csaba Szepesvári: Algorithms for Reinforcement Learning
- Dimitri P. Bertsekas: Dynamic Programming and Optimal Control
- Sheldon M. Ross: Introduction to Stochastic Dynamic Programming

### **Other resources:**

- I will provide additional reading materials during the course.
  - Start with: Review of Probability and Statistics.
- 

## **LEARNING GOALS AND OBJECTIVES**

The objective is to introduce you to the theory, techniques, and applications of Reinforcement Learning. This will necessitate some use of basic probability and statistics theory and computer manipulations of data. The goal is not that you will become experts on computer software, but rather to give you an overview on how Reinforcement Learning models and techniques are used

in practice. Hence, much of the material will be presented in a setting of practical situations. The focus is on ways of modeling and structural ideas about systems that evolve randomly in time.

Specific goals: The first part of the course will cover foundational material on Multi Armed Bandits. Then after a review of Markovian Decision Models (MDPs), we will focus on the problem of efficient exploration exploitation of a system of data collection for estimating policies that maximize the long run value. We will analyze methods such as UCB polices, optimistic programming, posterior sampling, and temporal difference learning. In addition, we will explore applications in Marketing, Inventory control, Health care and Marketing.

---

## PREREQUISITES

- a) Proficiency in Python and/or R.  
Class assignments will be in Python (using numpy) or R .
- b) Basic Probability and Statistics

---

## ACADEMIC INTEGRITY

*I do NOT tolerate cheating.* Students are responsible for understanding the RU Academic Integrity Policy (<http://academicintegrity.rutgers.edu/>). I will strongly enforce this Policy and pursue *all* violations. On all examinations and assignments, students must sign the RU Honor Pledge, which states, “On my honor, I have neither received nor given any unauthorized assistance on this examination or assignment.” [I will screen all written assignments through *SafeAssign* or *Turnitin*, plagiarism detection services that compare the work against a large database of past work.] Don’t let cheating or plagiarism destroy your hard-earned opportunity to learn and advance. See [business.rutgers.edu/ai](http://business.rutgers.edu/ai) for more details.

### Guidance on the use of AI at Rutgers

As noted in [Rutgers Academic Integrity Policy 10.2.13](#), the principles of academic integrity require that students make sure that all submitted coursework be “the student’s own and created without the aid of impermissible technologies, materials, or collaborations.”

---

## ATTENDANCE AND PREPARATION POLICY

Grade performance is a demonstrated function of attendance, preparation and participation. You can get behind very easily by skipping classes, resulting in a poor understanding of the material, which will show up as a poor grade for the class. Any class sessions missed by the student are the student’s responsibility to make up, not the instructor’s. Late arrival that causes disruption, early departure that causes disruption, excessive conversation among students (a disruption in its own right), inappropriate use of electronic devices that cause disruptions, and other actions that

disrupt the classroom are unacceptable.

If I am to be absent, my department chair or I will send you notice via email and Blackboard as far in advance as possible. If you are to be absent, report your absence in advance at <https://sims.rutgers.edu/ssra/>. If your absence is due to religious observance, a Rutgers-approved activity, illness, or family emergency/death and you seek makeup work, also send me an email with full details and supporting documentation within 3 days of your first absence.

If your absence is due to religious observance, a Rutgers-approved activity, illness, or family emergency/death and you seek makeup work, also send an email with full details and supporting documentation within 3 days of your absence. For weather emergencies, consult the campus home page. If the campus is open, class will be held. Expect me to arrive on time for each class session. I expect the same of you. Expect me to remain for the entirety of each class session. I expect the same of you. Expect me to prepare properly for each class session. I expect the same of you. Complete all background reading and assignments. You cannot learn if you are not prepared. The minimum expectation is that for each 3-hour class session, you have prepared by studying for at least twice as many hours. Expect me to participate fully in each class session. I expect the same of you. Stay focused and involved. You cannot learn if you are not paying attention.

---

## TESTS, PROJECTS AND GRADING

A combination of lectures, homework, and coding assignments, students will become well versed in key ideas and techniques for DP. Assignments will include the basics of reinforcement learning in addition, students will do and present an DP final project. Your grade<sup>1</sup> will be based on a final Project, homework assignments, and class participation, as follows:

- ◊ Mid Term Quiz 25%
- ◊ Final Project 50%
- ◊ Homework and class participation and conduct 25%

All exams will be take-home.

Your final grade is not subject to negotiation. If you feel I have made an error, submit your written argument to me within one week of receiving your final grade. Clarify the precise error I made and provide all due supporting documentation. If I have made an error, I will gladly correct it. But I will adjust grades only if I have made an error.

<sup>1</sup>Your final grade is not subject to negotiation. If you feel I have made an error, submit your written argument to me within one week of receiving your final grade. Clarify the precise error I made and provide all due supporting documentation. If I have made an error, I will gladly correct it. But I will adjust grades only if I have made an error.

---

## CLASSROOM CONDUCT

The course will be largely taught using Zoom presentation. Class-related material (lecture notes, messages, etc.) will be posted on the class site. Additional visual material and demos may be shown in some classes. Most of your work will take place outside the classroom, as you study, and apply the material to which you are introduced in class.

**Homework Assignments** are designed to help you learn the material discussed in class. In addition doing a thorough job on the homework assignments is the best preparation for the quiz and the final examination. There are three types of assignments: read, prepare, and hand in.

- **Read:** When the assignment is to read some material, this reading is an important introduction to the topics to be discussed in class. I will make the assumption that you have done the reading before class and have understood much (but not necessarily all) of it. When the assignment is to read a problem, that problem will often be used in class to introduce new concepts.

- **Prepare:** Fully analyze the problem. Be ready to discuss it in class using Zoom Share-Screen share screen , with the numbers computed, etc. I will call on people, so please be ready.

- **Hand In:** The same as prepare, but you must turn in your analysis. All written assignments must be handed in at the beginning of class on the day they are due, and so you will probably want to make a copy of your assignment for reference during class. All written assignments will be graded. These assignments should be submitted in typed form using a word processor. Please write your name, RUID and email on all homework submitted. Team work on this homework is not allowed. Unless a documented reason is produced for unusual circumstances, late submissions will not be accepted more than a week late.

Your class participation will be evaluated subjectively, but will rely upon measures of punctuality, attendance, familiarity with the required readings, relevance and insight reflected in classroom questions, and commentary. Your class participation will be judged by what you add to the class environment, regardless of your technical background. Although several lectures will be didactic, we will rely heavily upon interactive discussion within the class. Students will be expected to be familiar with the readings, even though they might not understand all of the material in advance. In general, questions and comments are encouraged. Comments should be limited to the important aspects of earlier points made, and reflect knowledge of the readings. You may be called on to answer questions about the homework or classroom discussion. Your classroom participation evaluation is based on the extent to which you contribute to the learning environment. However, correcting a mistake of the professor and asking what appear to be “dumb questions” about what is being covered are positive contributions. In the case of so-called “dumb questions,” very often half of the class will have the same questions in mind and are relieved to have them asked.

**Other requirements are:** On-time arrival to the Zoom classes, with uninterrupted attendance for the duration.

Maintenance of a professional atmosphere - use respectful comments and mute the microphone when not speaking.

Refraining from distracting or disrespectful activities, e.g., avoiding side conversations.

Courtesy towards all participants in the Zoom 'classroom'.

If your absence is due to religious observance, a Rutgers-approved activity, illness, or family emergency/death and you seek makeup work, also send an email with full details and supporting documentation within 3 days of your absence.

---

## **COURSE SCHEDULE**

Week 1. An overview of exploration/exploitation models. The Multi armed bandit (MAB) problem and Robbins early work.

◊ Week 2. Posterior Sampling and Bayes Methods for MABs.

◊ Week 3. Gittins indices and their computation.

◊ Week 4. UCB policies for MABs.

◊ Week 5. A review of optimization techniques for Markovian Decision Processes (MDPs).

◊ Week 6. Continuation of Week 5

◊ Week 7. UCB and e-greedy policies for MDPs, the pitfalls of 'Certainty equivalence'.

◊ Week 8. Optimistic Programming, Transformers and Deep Learning methods.

◊ Week 9. Continuation of Week 8

◊ Week 10. Continuation of Week 9

◊ Week 11. Applications to Health Care systems, and Marketing.

◊ Week 12. Review of the Vowpal Wabbit (VW) open source learning library.

◊ Week 13. Continuation of Week 12, Contextual bandits data visualization with Python.

◊ Week 14. End-to-load local loop for reinforcement learning.

---

## **SUPPORT SERVICES**

If you need accommodation for a *disability*, obtain a Letter of Accommodation from the Office of Disability Services. The Office of Disability Services at Rutgers, The State University of New

Jersey, provides student-centered and student-inclusive programming in compliance with the Americans with Disabilities Act of 1990, the Americans with Disabilities Act Amendments of 2008, Section 504 of the Rehabilitation Act of 1973, Section 508 of the Rehabilitation Act of 1998, and the New Jersey Law Against Discrimination. More information can be found at [ods.rutgers.edu](http://ods.rutgers.edu).

[Rutgers University-New Brunswick ODS phone (848)445-6800 or email [dsoffice@echo.rutgers.edu](mailto:dsoffice@echo.rutgers.edu)]

[Rutgers University-Newark ODS phone (973)353-5375 or email [ods@newark.rutgers.edu](mailto:ods@newark.rutgers.edu)]

If you are ***pregnant***, the Office of Title IX and ADA Compliance is available to assist with any concerns or potential accommodations related to pregnancy.

[Rutgers University-New Brunswick Title IX Coordinator phone (848)932-8200 or email [jackie.moran@rutgers.edu](mailto:jackie.moran@rutgers.edu)]

[Rutgers University-Newark Office of Title IX and ADA Compliance phone (973)353-1906 or email [TitleIX@newark.rutgers.edu](mailto:TitleIX@newark.rutgers.edu)]

If you seek ***religious accommodations***, the Office of the Dean of Students is available to verify absences for religious observance, as needed.

[Rutgers University-New Brunswick Dean of Students phone (848)932-2300 or email [deanofstudents@echo.rutgers.edu](mailto:deanofstudents@echo.rutgers.edu)]

[Rutgers University-Newark Dean of Students phone (973)353-5063 or email [DeanofStudents@newark.rutgers.edu](mailto:DeanofStudents@newark.rutgers.edu)]

If you have experienced any form of ***gender or sex-based discrimination or harassment***, including sexual assault, sexual harassment, relationship violence, or stalking, the Office for Violence Prevention and Victim Assistance provides help and support. More information can be found at <http://vpva.rutgers.edu/>.

[Rutgers University-New Brunswick incident report link:

<http://studentconduct.rutgers.edu/concern/>. You may contact the Office for Violence Prevention and Victim Assistance at (848)932-1181]

[Rutgers University-Newark incident report link:

[https://cm.maxient.com/reportingform.php?RutgersUniv&layout\\_id=7](https://cm.maxient.com/reportingform.php?RutgersUniv&layout_id=7) . You may also contact the Office of Title IX and ADA Compliance at (973)353-1906 or email at [TitleIX@newark.rutgers.edu](mailto:TitleIX@newark.rutgers.edu). If you wish to speak with a staff member who is confidential and does **not** have a reporting responsibility, you may contact the Office for Violence Prevention and Victim Assistance at (973)353-1918 or email [run.vpva@rutgers.edu](mailto:run.vpva@rutgers.edu)]

**Bias incidents:** an act – either verbal, written, physical, or psychological that threatens or harms a person or group on the basis of actual or perceived race, religion, color, sex, age, sexual orientation, gender identity or expression, national origin, ancestry, disability, marital status, civil union status, domestic partnership status, atypical heredity or cellular blood trait, military service or veteran status.

**Bias incidents can be reported online at:**

[New Brunswick Bias Incident Report Form](#)

[Newark Bias Incident Report Form](#)

If students who have experienced a temporary condition or injury that is adversely affecting their ability to fully participate, you should submit a request via <https://temporaryconditions.rutgers.edu>.

If you are a military **veteran** or are on active military duty, you can obtain support through the Office of Veteran and Military Programs and Services. <http://veterans.rutgers.edu/>

If you are in need of **mental health** services, please use our readily available services.  
[Rutgers University-Newark Counseling Center: <http://counseling.newark.rutgers.edu/>]  
[Rutgers Counseling and Psychological Services–New Brunswick: <http://rhscaps.rutgers.edu/>]

If you are in need of **physical health** services, please use our readily available services.  
[Rutgers Health Services – Newark: <http://health.newark.rutgers.edu/>]  
[Rutgers Health Services – New Brunswick: <http://health.rutgers.edu/>]

If you are in need of **legal** services, please use our readily available services:  
<http://rusls.rutgers.edu/>

Students experiencing difficulty in courses due to **English as a second language (ESL)** should contact the Program in American Language Studies for supports.  
[Rutgers–Newark: [PALS@newark.rutgers.edu](mailto:PALS@newark.rutgers.edu)]  
[Rutgers–New Brunswick: [eslpals@english.rutgers.edu](mailto:eslpals@english.rutgers.edu)]

If you are in need of additional **academic assistance**, please use our readily available services.  
[Rutgers University-Newark Learning Center: <http://www.ncas.rutgers.edu/rlc>]  
[Rutgers University-Newark Writing Center: <http://www.ncas.rutgers.edu/writingcenter>]  
[Rutgers University-New Brunswick Learning Center: <https://rlc.rutgers.edu/>]

[Optional items that many faculty include:

- Students must sign, date, and return a statement declaring that they understand the RU Academic Integrity Policy.
- Students must sign, date, and return a statement declaring that they understand this syllabus.]

---

## CODE OF PROFESSIONAL CONDUCT

[If you prefer to direct students to the conduct policy online instead, please use the following link and place it beneath the header above:

<https://myrbs.business.rutgers.edu/students/code-professional-conduct>]

Rutgers Business School is recognized for its high-quality education. To that end, maintaining the caliber of classroom excellence, whether in person or online, requires students to adhere to the same behaviors expected in professional career environments. These include the following principles:

## **Discussion and Correspondence**

- Each student is encouraged to participate actively in class discussions and exercises. Substantive dialogue requires a degree of mutual respect, willingness to listen, and tolerance of opposing points of view. Disagreement and the challenging of ideas must happen in a supportive and sensitive manner. Hostility and disrespectful behavior will not be tolerated.
- In correspondence and in the classroom, students should demonstrate respect in how they address instructors. Students should use proper titles unless there is an explicit understanding that the instructor accepts less formal alternatives. Similarly, appropriate formatting in electronic communication and timely responsiveness are all expectations in every professional interaction, including with instructors. Everything said and written should demonstrate respect and goodwill.

## **Punctuality and Disruption**

- Class starts and ends promptly at the assigned periods. Students are expected to be in their seats or present online and ready to begin class on time.
  - Take your responsibility to attend class seriously. Your attendance is a critical element of the learning experience for in-person classes. Failure to show up disrupts your learning and signals disrespect to your peers and instructors. (Of course, illness is a legitimate exception requiring advanced reporting to the [University](#) and your instructors.)
  - Barring emergencies and within reason, students are expected to remain in their seats for the class duration. In person, packing belongings before the end of class disturbs both other students and the instructor. Online, attending to other tasks is distracting. In addition, even if webcams are not required in your course, your attention is fundamentally lacking if you are engaged in multiple tasks simultaneously.

## **Technology**

- The use of technology is sanctioned only as permitted by the course instructor. As research on learning shows, peripheral use of technology in classes negatively impacts the learning environment in three ways:
  1. Individual learning and performance directly suffer, resulting in the systemic lowering of grades earned.
  2. In the classroom, one student's use of technology automatically diverts and captures other people's attention, thus impeding their learning and performance. Moreover, even minor infractions have a spillover effect and result in others doing the same.
  3. Subverting this policy (e.g., using a phone during class, even if hidden below the table or out of sight from your webcam; tapping on a smartwatch; using a laptop

for non-course related matters) is evident to the course instructor and offensive to the principles of decorum in a learning environment.

- Networking, computing, and associated resources in the trading rooms, advanced technology rooms, and general classrooms are to be used in the manner intended.
- Sharing links to private online classes, attempting to join an online class you are not enrolled in, or posting disruptive content during these sessions are strictly prohibited and may lead to disciplinary action.
- For more instructions on information technology resources at Rutgers University, please refer to the [Acceptable Use Policy for Information Technology Resources](#).

### **Misappropriating Intellectual Property**

- Almost all original work is the intellectual property of its authors. These works may include syllabi, lecture slides, recorded lectures, homework problems, exams, and other materials, in either printed or electronic form. The authors may hold copyrights in these works, which U.S. statutes protect. Copying this work or posting it online (on sites such as Chegg or Course Hero) without the author's permission may violate the author's rights. More importantly, these works are the product of the author's efforts; respect for these efforts and the author's intellectual property rights are important values that members of the university community take seriously.
- For more instructions on copyright protections at Rutgers University, please refer to the [Rutgers Libraries](#).

Rutgers Business School is committed to the highest standards of integrity. We value mutual respect and responsibility, as these are fundamental to our educational excellence inside and outside the classroom.