Northwestern | SCHOOL OF PROFESSIONAL STUDIES

Master of Science in Data Science (MSDS)

Course Syllabus: Spring 2019

INTELLIGENT SYSTEMS AND ROBOTICS (ISAR)
MSDS 464 in AI Specialization

Instructor: Shreenidhi Bharadwaj
Email: shreenidhi.bharadwaj@northwestern.edu
LinkedIn: https://www.linkedin.com/in/ssbharadwaj/
Emails (QnA): Response within 24-48 hours

TA: Christopher Fiore
Email: ChristopherFiore2015@u.northwestern.edu

Sync Sessions:
Mondays: 7PM – 8:30PM CST

COURSE DESCRIPTION

This course introduces reinforcement learning as an approach to intelligent systems, emphasizing applications such as robotic processes automation, conversational agents and robotics that mimic human behavior. Students implement intelligent agents to solve both discrete- and continuous-valued sequential decision-making problems. Students develop, debug, train, visualize, and customize programs in a variety of learning environments. The course reviews Markov decision processes, dynamic programming, temporal difference learning, Monte Carlo reinforcement learning, eligibility traces, the role of function approximation, and the integration of learning and planning. This is a case study and project-based course with a substantial programming component.

Recommended prior course: MSDS 458-DL Artificial Intelligence and Deep Learning.
Prerequisites:
(1) MSDS 420-DL Database Systems and Data Preparation or CIS 417 Database Systems Design and Implementation &
(2) MSDS 422-DL Practical Machine Learning or CIS 435 Practical Data Science Using Machine Learning.

COURSE OBJECTIVES

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

- Identify the use cases and business applications for reinforcement learning
- Differentiate between supervised, unsupervised learning from reinforcement learning.
- Describe Intelligent Agents and be able to demonstrate the workings of it.
- Interact with a conversational agent and explain the workings of an interactive chatbot
- Define the key features of reinforcement learning that distinguishes it from AI and non-interactive machine learning.
- Implement RL algorithms such as a deep RL algorithm, including imitation learning
- Given an application problem (e.g. from computer vision, robotics, etc) be able to determine if it should be formulated as a RL problem and state what algorithms are best suited for addressing it.
- Describe (list and define) multiple criteria for analyzing RL algorithms and evaluate algorithms on these metrics: e.g. regret, sample complexity, computational complexity, empirical performance, convergence, etc.
- Compare at least two approaches for addressing the exploration vs exploitation challenge in terms of performance, scalability, complexity of implementation, and theoretical guarantees.

COURSE PREREQUISITES

Following are the prerequisites to be successful in the course.

1. MSDS 420-DL Database Systems and Data Preparation
2. MSDS 422-DL Practical Machine Learning
3. MSDS 458 Artificial Intelligence and Deep Learning

TEXTBOOKS

Primary Textbooks (Required)

- Sutton, R., & Barto, A. (2018). Reinforcement learning: An introduction (Second ed., Adaptive computation and machine learning). Cambridge, Massachusetts: The MIT Press.
    - References
        - Author's website with resources including PDF copy of book
        - Python Implementation of Reinforcement Learning: An Introduction

- Lapan, M., & Safari, an O'Reilly Media Company. (2018). Deep Reinforcement Learning Hands-On (1st ed.). Packt Publishing.
    - References
        - Hands-on Deep Reinforcement Learning published by Packt

Optional Readings and Resources (Reference Textbooks)

- Algorithms for Reinforcement Learning - Csaba Szepesvári
    - References
        - https://sites.ualberta.ca/~szepesva/RLBook.html

- Géron, A., 2017. Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. Sebastopol Calif.: O'Reilly. [ISBN-13: 978-1491962299]
    - References
        - https://github.com/ageron/handson-ml

- Goodfellow, I., Bengio, Y., and Courville, A., 2016. Deep Learning. Cambridge, Mass.: MIT Press. [ISBN-13: 978-0262035613]
    - References
        - https://mitpress.mit.edu/books/deeplearning

Note: This text is highly-regarded. It is also very mathematical and abstract. It is useful to know about this as a classic reference, but other texts are more suitable as introductions.

NOTE:

- Because the Intelligent Agents & Robotics are fast-moving fields, new readings (required and optional) may be introduced throughout the term.
- Additional Required Readings will be identified within specific Modules and will typically be classical and recent papers in artificial intelligence, neural networks, and deep learning. In all cases, web links to accessible sources will be provided.
- Assigned readings posted on Canvas, including timely news articles and academic research that you will read to complete some assignments and participate in discussion forums.

COURSE RESERVES

Selected readings are available through the Course Reserves. For assistance with Course Reserves, use electronic mail: e-reserve@northwestern.edu. To ask a librarian for assistance, visit Northwestern's Ask A Librarian page.

ASSIGNMENTS & GRADING SCALE

Grading: Grading and feedback turnaround will be one week from the due date. You will be notified if turnaround will be longer than one week. The discussion forums and written assignments will be graded based on specific grading guidelines.
Assignments: There are no quizzes or exams. Rubrics are provided for assignments.

Class Participation	Graded discussion threads (10 points per week)	100 Points
Assignment 1	First Research/Programming Assignment (Week 3)	100 Points
Assignment 2	Second Research/Programming Assignment (Week 5)	100 Points
Assignment 3 Proposal	Proposal for Fourth Research /Programming Assignment (Week 6)	0 Points
Assignment 4	Third Research/Programming Assignment (Week 7)	100 Points
Assignment 5	Fourth Research/Programming Assignment (Week 10)	100 Points
	Total	500 Points

Grade	Percentage	Total Points (out of 500)
A	93%–100%	465– 500 points
A-	90%–92%	450 – 464 points
B+	87%–89%	435 – 449 points

B	83%–86%	415 – 434 points
B-	80%–82%	400 – 414 points
C+	77%–79%	385 – 399 points
C	73%–76%	365 – 384 points
C-	70%–72%	350 – 364 points
F	0%–69%	0 – 349 points

COURSE PROCEDURES

Assignment Due Dates

Each Monday introduces a new week, and its corresponding weekly Module. All discussions and assignments are due Sunday evening at 11:55 p.m. Central time.

Participation and Attendance

This course follows the asynchronous distance learning approach of Northwestern University School of Profession Studies. The course does not meet at a particular time each week. Learning objectives and assessments are supported through classroom elements that can be accessed at any time. To measure class participation (or attendance), your participation in threaded discussion boards is required, graded, and paramount to your success in this course.

Real-time, synchronous meetings (Sync Sessions) are scheduled for Thursday evenings on selected weeks. Sync Sessions are conducted with WebEx. While your attendance is highly encouraged, it is not required. You will not be graded on your attendance or participation. All Sync Sessions are recorded.

Refer to the weekly schedule at the end of this syllabus for details about weekly learning objectives, required and optional readings, assignments, and Sync Sessions.

Late Assignment Turn-Ins

Students should provide written notification of late assignment work 24 hours prior to the deadline. A grace day is allowed for those who provide late work notification. Late papers may be subject to point reductions.

Discussion Forums

The purpose of the discussion boards is to allow students to freely exchange ideas. It is imperative to remain respectful of all viewpoints and positions and, when necessary, agree to respectfully disagree. While active and frequent participation is encouraged, cluttering a discussion board with inappropriate, irrelevant, or insignificant material will not earn additional points and may result in receiving less than full credit. Frequency matters but contributing content that adds value is paramount. Please remember to cite all sources—when relevant—in order to avoid plagiarism. Please post your viewpoints first and then discuss others' viewpoints.

The quality of your posts and how others view and respond to them are the most valued. A single statement mostly implying "I agree" or "I do not agree" is not counted as a post. Explain, clarify, politely ask for details, provide details, persuade, and enrich communications for a great discussion experience. Please note, there is a requirement to respond to at least two fellow class members posts. Also, remember to cite all sources—when relevant—in order to avoid plagiarism.

Online Communication Etiquette

Beyond interacting with your instructor and peers in discussions, you will be expected to communicate by Canvas message, email, and sync session. Your instructor may also make themselves available by phone or text. In all contexts, keep your communication professional and respect the instructor's posted availability. To learn more about professional communication, please review the Communicating Effectively with Faculty guide.

Just as you expect a response when you send a message to your instructor, please respond promptly when your instructor contacts you. Your instructor will expect a response within two business days. This will require that you log into the course site regularly and set up your notifications to inform you when the instructor posts an announcement, provides feedback on work, or sends you a Canvas message. For guidance on setting your notifications, please review How do I set my Canvas notification settings as a student? It is also recommended that you check your u.northwestern e-mail account regularly, or forward your u.northwestern e-mail to an account you check frequently.

In this class, in all interactions, it is important that we each remain respectful of the viewpoints and positions of others. When necessary (as some conversations may become spirited), we may have to "agree to disagree."
Discussion board posting should be used exclusively for topic-related postings; posting inappropriate, irrelevant, or insignificant material will actually earn you a negative score.

Frequency is not unimportant, but content of the message is paramount. A rubric for evaluating discussion posts will be provided within the first week of class. When you use material from an external source, please follow full professional (academic) citation styles, so that you (and others) can reference this material later. (Keep in mind that citing sources is a means of avoiding plagiarism.) For more information, read the 10 Rules of Netiquette.

Study Teams (Required)

Student study teams is utilized in this course as a means to foster a collaborative learning environment. Blue Jeans is available as a conferencing tool. Preliminary groups have/can been/be set up under the People/Study Teams tab of Canvas. Each student is encouraged to join a study team. It may make sense to join a team based on time zone (Eastern, Central, or Mountain/Pacific) and preferred personal computer operating system (Mac OSX or Windows). It is recommended that teams consist of no more than four students.

STUDENT SUPPORT SERVICES

AccessibleNU

This course is designed to be welcoming to, accessible to, and usable by everyone, including students who are English-language learners, have a variety of learning styles, have disabilities, or are new to online learning. Be sure to let me know immediately if you encounter a required element or resource in the course that is not accessible to you. Also, let me know of changes I can make to the course so that it is more welcoming to, accessible to, or usable by students who take this course in the future.

Northwestern University and AccessibleNU are committed to providing a supportive and challenging environment for all undergraduate, graduate, professional school, and professional studies students with disabilities who attend the University. Additionally, the University and AccessibleNU work to provide students with disabilities and other conditions requiring accommodation a learning and community environment that affords them full participation, equal access, and reasonable accommodation. The majority of accommodations, services, and auxiliary aids provided to eligible students are coordinated by AccessibleNU, which is part of the Dean of Students Office.

SPS Student Services

The Department of Student Services supports the academic and professional growth of SPS students. The Student Services team guides students through academic planning, policies, and administrative procedures, and promotes a supportive environment to foster student success. Students are encouraged to actively make use of the resources and staff available to assist them: Academic and Career Advisers, Counseling and Health Services, Student Affairs, Legal Services, Financial Aid and Student Accounts, among other services.

For a comprehensive overview of course and program processes and policies and helpful student resources, please refer to your SPS Student Handbook.

ACADEMIC SUPPORT SERVICES

Northwestern University Library

As one of the leading private research libraries in the United States, Northwestern University Library serves the educational and information needs of its students and faculty as well as scholars around the world. Visit the Library About page for more information or contact Distance Learning Librarian Tracy Coyne at 312-503-6617 or tracy-coyne@northwestern.edu.

Program-Specific Library Guides

- Information Systems
- Data Science

Additional Library Resources

- Connectivity: Campus Wireless and Off-Campus Access to Electronic Resources

- Reserve a Library Study Room
- Sign up for an in-person or online Research Consultation Appointment
- Getting Available Items: Delivery to Long-Distance Patrons
- Social Science Data Resources
- Resources for Data Analysis

The Writing Place

The Writing Place is Northwestern's center for peer writing consultations. Consultations are free and available to anyone in the Northwestern community: undergraduates, graduate students, faculty, or staff. To book an appointment, go to The Writing Place website.

The Math Place

The Math Place is a free tutorial service provided to students currently enrolled in Northwestern University's School of Professional Studies courses or in other Northwestern University courses. Students of all levels can benefit from the individual tutoring provided from this service, whether they are taking undergraduate or graduate level courses. To book an appointment, go to The Math Place website.

Academic Integrity at Northwestern

Students are required to comply with University regulations regarding academic integrity. If you are in doubt about what constitutes academic dishonesty, speak with your instructor or graduate coordinator before the assignment is due and/or examine the University Web site. Academic dishonesty includes, but is not limited to, cheating on an exam, obtaining an unfair advantage, and plagiarism (e.g., using material from readings without citing or copying another student's paper). Failure to maintain academic integrity will result in a grade sanction, possibly as severe as failing and being required to retake the course and could lead to a suspension or expulsion from the program. Further penalties may apply. For more information, visit The Office of the Provost's Academic Integrity page.

Some assignments in SPS courses may be required to be submitted through Turnitin, a plagiarism detection and education tool. You can find an explanation of the tool here.

COURSE TECHNOLOGY REQUIREMENTS

This course will involve a number of different types of interactions. These interactions will take place primarily through the Canvas system. Please take the time to navigate through the course and become familiar with the course syllabus, structure, and content and review the list of resources below.

Canvas

The Canvas Student Center includes information on communicating in Canvas, navigating a Canvas course, grades, additional help, and more. The Canvas at Northwestern website provides information of getting to know Canvas at Northwestern and getting Canvas support. The Canvas Student Guide

provides tutorials on all the features of Canvas. For additional Canvas help and support, you can always click the Help icon in the lower left corner to begin a live chat with Canvas support or contact the Canvas Support Hotline. The Canvas Accessibility Statement and Canvas Privacy Policy are also available.

Software & Technology Resources

Required software for this course is open source and freely available on the web. Python is the primary programming environment. Personal computer software is available for Mac OS X and Windows systems. Students will also have access to the Data Science Computing Cluster from the School of Professional Studies. Additional information about software solutions will be provided on the Canvas course site under Modules / Technical Resources.

(Optional) Lynda.com offers a wide range of technical training materials, and Northwestern University has a contract with Lynda.com to provide these courses at no cost to registered students. The program is described at: http://www.northwestern.edu/hr/workplace-learning/lynda

Of special interest are Lynda.com courses in Python programming, JavaScript, web development, and Git/GitHub. The Lynda.com Learning Path *Become a Front-End Developer* includes a one-hour and twenty-minute lecture entitled *Up and Running with Git and GitHub.* Another Lynda.com course *Git Essential Training* is a more complete six-hour course on Git technology. Git runs offline on personal computers, and GitHub provides a server-based storage facility for individual assignments and team projects. Repositories for data, code, and reports can be shared among team members.

This course will use UIPath tool to get a hands-on understanding of the Robotic Process Automation, Python as the primary programming language along with TensorFlow, Keras, PyTorch & OpenAI Gym frameworks. Python can be downloaded at Anaconda. Download and install the Python 3.6 version of Anaconda. Additional documentation on downloading and installing Anaconda can be found at Anaconda Documentation. After installation, you should be able to access the Anaconda Navigator. To work with a Python-aware editor and an iPython shell, Spyder is recommended. This graphical user interface can be launched via the Anaconda Navigator. This course will make use of Jupyter notebooks extensively to build and store Python files for various assignments. Jupyter can also be launched via the Anaconda Navigator. Additional details relating to the above are listed below.

- UI Path
  - https://www.uipath.com/freetrial-or-community
  - https://www.uipath.com/developers/community-edition/license-agreement

- Anaconda
  - https://www.anaconda.com/download/
  - https://docs.anaconda.com/anaconda/eula/
- OpenCV
  - https://opencv.org/
  - https://opencv.org/license.html
- JetBrains (Pycharm)
  - https://www.jetbrains.com/student/
  - https://www.jetbrains.com/legal/agreements/user.html
- TensorFlow
  - https://www.tensorflow.org/
  - https://github.com/tensorflow/tensorflow/blob/master/LICENSE
- Keras
  - https://keras.io/
  - https://github.com/keras-team/keras/blob/master/LICENSE
- OpenAI
  - https://gym.openai.com/
  - https://github.com/openai/gym/blob/master/LICENSE.md
- PyTorch
  - https://pytorch.org/get-started/locally/
  - https://github.com/pytorch/vision/blob/master/LICENSE

Zoom and Blue Jeans

We will use Zoom for optional synchronous meetings (sync sessions). You can review the Privacy Policy here and the Accessibility statement here.

Please note that any scheduled synchronous meetings are optional. While your attendance is highly encouraged, it is not required, and you will not be graded on your attendance or participation. These synchronous sessions will be recorded, so you will be able to review the session afterwards.

Blue Jeans video conferencing is available for student use in recommended study teams.

Panopto

Videos in this course may be hosted in Panopto. If you have not used Panopto in the past, you may be prompted to login to Panopto for the first time and authorize Panopto to access your Canvas account. You can learn more about using Panopto and login to Panopto directly by visiting the Panopto guide on the Northwestern IT Resource Hub. Depending on the assignment requirements of this course, you may be asked to create videos using Panopto in addition to viewing content that your instructor has provided through Panopto.

The Panopto Privacy Policy and the Accessibility Features on Panopto are also available.

Minimum Required Technical Skills

Students in an online program should be able to do the following:

- Communicate via email and Canvas discussion forums.
- Use web browsers and navigate the World Wide Web.
- Use the learning management system Canvas.
- Use integrated Canvas tools (e.g., WebEx, Bluejeans, Panopto, Course Reserves).
- Use applications to create documents and presentations (e.g., Microsoft Word, PowerPoint).
- Use applications to share files (e.g., Box, Google Drive).
- Use software for predictive analytics (e.g., Python, TensorFlow, Keras).

Systems Requirements for Distance Learning

Students and faculty enrolled in SPS online master's degree programs should have access to a computer with Minimum System Requirements.

Technical Help and Support

The SPS Help Desk is available for Faculty, Students and Staff to support their daily IT needs. For additional technical support, contact the Northwestern IT Support Center.

COURSE CALENDAR

1) The pages that follow show the plan for topics and assignments week-by-week.
2) In the syllabus and Canvas course site, "week," "session," and "module" are synonymous.
3) Changes to the syllabus will be posted to the Canvas course site with an Announcement.

Weekly Calendar

- Week 1: Introduction to Intelligent Agents, Robotics – Applications & Frameworks
- Week 2: Introduction to Reinforcement Learning
- Week 3: Markov Decision Processes, First Research/Programming Assignment
- Week 4: Deep Q-Learning, Double Q-Learning, Replay Memory
- Week 5: Planning by Dynamic Programming, Second Research/Programming Assignment
- Week 6: Monte Carlo learning, Proposal for Fourth Research/Programming Assignment
- Week 7: Temporal difference learning, Third Research/Programming Assignment
- Week 8: Value Function Approximations, Deep Imitation Learning
- Week 9: Policy Gradient Methods, Exploration and Exploitation
- Week 10. Finalize and Submit Fourth Research/Programming Assignment

Week 1: Introduction to Intelligent Systems and Robotics

In this module we will introduce the topics relating to Intelligent Agents, Software Robotics and Robotic Process Automation (RPA). We will use flowcharts to document enterprise processes spanning multiple technologies and system boundaries while learning automation best practices. You will learn to create software robots & automate tasks while discovering the value-add to the enterprise

Learning Objectives

- Describe key terms & concepts relating to Intelligent Agents & Robotics.
- Describe Robotic Process Automation (RPA) and be able to demonstrate the workings of it.
- Identify how Robotic Process Automation fits into Artificial Intelligence domain
- Explain the Optical Character Recognition (OCR), web scraping technologies and its workings within RPA.
- Create flowcharts and automate tasks using Robotic Processes.

Textbook Readings (Required)

- None

Resources (Optional)

- https://www.cio.com/article/3236451/what-is-rpa-robotic-process-automation-explained.html
- https://www.uipath.com/developers/video-tutorials/introduction-to-uipath
- https://www.youtube.com/watch?v=LyEOo8AtGLQ
- https://www.uipath.com/video/uipath-abbyy-framework-demo
- https://robot.uipath.com/docs/introduction
- https://studio.uipath.com/docs/about-data-scraping

Discussion (10 points)

What are some of the business processes that can be automated via RPA in the HealthCare, Retail or finance domain? Provide 2 use cases with concrete examples.

Sync Sessions

Mini-Lecture: Course & module 1 Overview, hands on session relating to RPA (UIPath)
Sync Session attendance is optional. Each Sync Session is recorded.

Week 2: Introduction to Reinforcement Learning (RL)

In this module we will introduce the branches of machine learning, Reinforcement Learning (RL) and the characteristics along with examples of RL. We will incorporate a time dimension into learning equations which will get us closer to the human perception of artificial Intelligence.

Learning Objectives

- Describe how RL is related to and differs from other ML disciplines (supervised & unsupervised learning)
- Describe the theoretical foundations of Reinforcement Learning along with the main RL formalisms and how they are related to one another
- Identify Components of a Reinforcement Learning Agent and be able to categorize them

Textbook Readings (Required)

- Sutton & Barto, Chapters 1,2
- Max Laplan Chapters 1,2,3
- TensorFlow and deep reinforcement learning, without a PhD (Google I/O 18)
    - https://www.youtube.com/watch?v=t1A3NTttvBA

Resources (Optional)

- Géron Chapter 16
- https://blog.openai.com/learning-dexterity/
- https://www.youtube.com/watch?v=t1A3NTttvBA
- https://www.oreilly.com/learning/how-to-build-a-robot-that-sees-with-100-and-tensorflow
- https://towardsdatascience.com/reinforcement-learning-demystified-36c39c11ec14

Discussion (10 points)

Discussion: What are some of the examples/use cases where reinforcement learning is applicable.

Sync Sessions

Mini-Lecture: Course Overview

Sync Session attendance is optional. Each Sync Session is recorded

Week 3: Markov Decision Processes

In this module we will introduce and walk through the Markov Processes, Markov Reward Processes, Markov Decision Processes & extensions to MDPs. We will be using MDP to describe an environment for reinforcement learning, where the environment is fully observable since most RL problems can be formalized as MDPs.

Learning Objectives

- Describe an environment for reinforcement learning (Agent-Environment)
- Describe MDPs (Markov Decision Processes) & interpret transition diagrams
- Differentiate between Markov Property, Markov Process, Markov Reward Process & Markov decision Process
- Analyze and demonstrate Tabular Solution Methods & Multi-armed Bandits Problem
- Describe the Bellman Equation of Optimality with examples

Textbook Readings (Required)

- Sutton & Barto, Chapter 3
- Max Laplan Chapters 4,5,13

Resources (Optional)

- https://towardsdatascience.com/reinforcement-learning-demystified-markov-decision-processes-part-1-bf00dda41690
- https://harderchoices.com/2018/02/11/finite-markov-decision-process-a-high-level-introduction/

Discussion (10 points)

What is Markov Property, Markov process & the Markov Reward Process? How are they tied up to the Markov Decision Process?

Sync Sessions

Mini-Lecture: Browser Automation & RL

Sync Session attendance is optional. Each Sync Session is recorded

Week 4: Deep Q-Learning, Double Q-Learning, Replay Memory

In this module we will understand the workings of DQN algorithm, which combines Q-learning with a deep neural network. Since DQN algorithm suffers from substantial overestimations in some games we will explore the idea behind the Double Q-learning algorithm, which can be generalized to work with large-scale function approximation. With specific adaptation to the DQN algorithm we will learn that the resulting algorithm not only reduces the observed overestimations, but also leads to much better performance on several games

Learning Objectives

- Demonstrate how we can make Q-learning work with deep networks
- Describe a generalized view of Q-learning algorithms & tricks for improving Q-learning in practice
- Implement Q-learning and understand how to extend Q-learning to continuous actions
- Analyze and be able to describe Double Deep Q-Learning with examples

Textbook Readings (Required)

- Max Laplan Chapters 6,7,8

Resources (Optional)

- https://deepmind.com/research/dqn/
- https://arxiv.org/abs/1509.06461
- https://www.youtube.com/watch?v=t1A3NTttvBA

Discussion (10 points)

What are the complications that arise from approximation of Q-Values with neural networks? What are some of the tricks for DQNs to improve their training stability & convergence?

Sync Sessions

Mini-Lecture: Stock Trading Using RL

Sync Session attendance is optional. Each Sync Session is recorded

Week 5: Dynamic Programming

In this module we will focus on the planning problems and go through a process of solving known MDPs using policy evaluation, policy iteration and value iteration. Additionally, given the complete model and specifications of the environment (MDP), we will walkthrough an example to evaluate optimal policies using Dynamic Programming for the agent to follow.

Learning Objectives

- Evaluate the differences between Policy Evaluation and Policy
- Improve on how the processes interact with each other and describe the Limitations of Dynamic Programming Approaches
- Describe and implement the Policy Iteration Algorithm
- Describe and develop the Value Iteration Algorithm

Textbook Readings (Required)

- Sutton & Barto, Chapters 4
- Max Laplan Chapters 12

Resources (Optional)

- https://www.analyticsvidhya.com/blog/2018/09/reinforcement-learning-model-based-planning-dynamic-programming/

Discussion / Assignment (10 points)

What is a goal oriented chatbot? What are some of the use cases and their real world applications?

Sync Sessions

Mini-Lecture: Chat Bots Training with RL

Sync Session attendance is optional. Each Sync Session is recorded

Week 6: Monte Carlo learning, Monte Carlo Tree Search

In this module we will go Monte Carlo (MC) methods that can learn directly from experience collected by interacting with the environment. We will walkthrough an example of MCTS that figures out the best move out of a set of moves by Selecting → Expanding → Simulating → Updating the nodes in tree to find the final solution.

Learning Objectives

- Describe and identify the differences between Prediction and Control
- Know how to use the MC method for predicting state values and state-action values
- Analyze the on-policy first-visit & off-policy MC control algorithm
- Explain benefits of MC algorithms over the Dynamic Programming approach

Textbook Readings (Required)

- Sutton & Barto, Chapters 5, 8

Resources (Optional)

- https://towardsdatascience.com/monte-carlo-tree-search-158a917a8baa

Discussion (10 points)

What are some of the examples/use cases where MCTS is applicable.

Sync Sessions

Mini-Lecture: Board Game of AlphaGo

Sync Session attendance is optional. Each Sync Session is recorded

Week 7: Temporal difference learning

In this module we will go through a process of understanding and implementing TD-Learning. TD-Learning is a combination of Monte Carlo and Dynamic Programming ideas. Like Monte Carlo, TD works based on samples and doesn't require a model of the environment. Like Dynamic Programming, TD uses bootstrapping to make updates

Learning Objectives

- Implement TD (0) for prediction and SARSA for on-policy control
- Describe the benefits of TD algorithms over MC and DP approaches
- Identify how n-step methods unify MC and TD approaches
- Evaluate and differentiate the backward and forward view of TD-Lambda

Textbook Readings (Required)

- Sutton & Barto, Chapters 6

Resources (Optional)

- https://medium.com/@violante.andre/simple-reinforcement-learning-temporal-difference-learning-e883ea0d65b0
- https://www.ias.informatik.tu-darmstadt.de/uploads/Teaching/AutonomousLearningSystems/Kunz_ALS_2013.pdf

Discussion (10 points)

TDL are general methods for learning to make long-term predictions about dynamic systems. Can you describe an application and how some of these methods are applicable?

Sync Sessions

Mini-Lecture

Sync Session attendance is optional. Each Sync Session is recorded

Week 8: Value Function Approximations, Imitation Learning

In this module we will go through function approximation and understand how it can be a solution for large MDPs. We will also get an understanding of imitation learning techniques that attempts to enable effective "programming by demonstration" to automate tasks, that people can demonstrate but find difficult to hand program

Learning Objectives

- Describe the motivation for Function Approximation over Table Lookup
- Analyze how to incorporate function approximation into existing algorithms
- Describe the convergence properties of function approximators and RL algorithms
- Apply batching using experience replay

Textbook Readings (Required)

- Sutton & Barto, Chapter 6, 9, 10
- Max Laplan Chapters 6,7,8

Resources (Optional)

- https://arxiv.org/abs/1807.09205
- https://www.disneyresearch.com/publication/data-driven-ghosting/
- http://karpathy.github.io/2016/05/31/rl/
- https://arxiv.org/abs/1607.05241
- https://arxiv.org/abs/1604.07316

Discussion / Assignment (10 points)

What is data driven ghosting and what are some of the real world use cases where it can be applied?

Sync Sessions

Mini-Lecture

Sync Session attendance is optional. Each Sync Session is recorded

Week 9: Policy Gradient Methods, Exploration and Exploitation

In this module we will work on understanding why instead of parameterizing the value function and doing greedy policy improvement we parameterize the policy and do gradient descent into a direction that improves it

Learning Objectives

- Analyze differences between value-based and policy-based Reinforcement Learning
- Implement and evaluate the REINFORCE Algorithm (Monte Carlo Policy Gradient)
- Describe Actor-Critic (AC) algorithms & Advantage Functions

Textbook Readings (Required)

- Sutton & Barto, Chapter 7,13
- Max Laplan Chapters 9,10

Resources (Optional)

- http://karpathy.github.io/2016/05/31/rl/
- https://arxiv.org/abs/1607.05241

Discussion / Assignment (10 points)

What is the difference between DQNs and actor critic methods? Can you provide applications where actor-critic algorithm can be used?

Sync Sessions

Mini-Lecture

Sync Session attendance is optional. Each Sync Session is recorded

Week 10: Final Assignment (Team Project)

During this week, students work in teams to submit their final project for evaluation.

Learning Objectives

No new learning objectives. This week is for final project and course feedback.

Textbook Readings (Required)

- None

Resources (Optional)

- None

Discussion / Assignment (10 points)

What have you learned from this course? What portions of the course were most beneficial to you and your career? Provide at least two suggestions for improving the course.

Sync Sessions

NA

Reference Materials

- A Brief Survey of Deep Reinforcement Learning, Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, Anil Anthony Bharath (https://arxiv.org/pdf/1708.05866.pdf)
- Deep Reinforcement Learning UC Berkeley. "CS294-112." CS 294-112. Accessed March 12, 2019. http://rail.eecs.berkeley.edu/deeprlcourse/.
- Reinforcement Learning from Stanford: CS234: Reinforcement Learning Winter 2019. Accessed March 15, 2019. http://web.stanford.edu/class/cs234/index.html
- Key Papers in Deep RL: https://spinningup.openai.com/en/latest/spinningup/keypapers.html
- "Key Papers in Deep RL." Key Papers in Deep RL - Spinning Up Documentation. Accessed March 12, 2019. https://spinningup.openai.com/en/latest/spinningup/keypapers.html
- Deep Learning Reading List: http://deeplearning.net/reading-list/
- "Reading List." Deep Learning. Accessed March 12, 2019. http://deeplearning.net/reading-list/
- Denny Britz."Reinforcement-learning." GitHub. March 02, 2019. Accessed March 12, 2019. https://github.com/dennybritz/reinforcement-learning
- Reinforcement Learning McGill: https://www.cs.mcgill.ca/~dprecup/courses/RL/lectures.html
- https://github.com/songrotek/Deep-Learning-Papers-Reading-Roadmap - a very good place to start, begins with the text that we'll be using; this is a curated list.
- Floodsung. "Floodsung/Deep-Learning-Papers-Reading-Roadmap." GitHub. March 16, 2018. Accessed March 15, 2019. https://github.com/floodsung/Deep-Learning-Papers-Reading-Roadmap .
- Papers with Code: https://github.com/zziz/pwc
- Zziz. "Zziz/pwc: Papers with Code." GitHub. January 06, 2019. Accessed March 15, 2019. https://github.com/zziz/pwc .
- Matrix Calculus for Deep Learning: https://explained.ai/matrix-calculus/index.html
- "The Matrix Calculus You Need for Deep Learning." The Matrix Calculus You Need for Deep Learning. Accessed March 15, 2019. https://explained.ai/matrix-calculus/index.html
- Deng, L. and Yu, D., Deep Learning: Methods and Applications: (a 197-page monograph): http://research.microsoft.com/pubs/209355/DeepLearning-NowPublishingVol7-SIG-039.pdf
- Deng, Li, and Dong Yu. "Deep Learning: Methods and Applications." Microsoft Research. May 01, 2014. Accessed March 15, 2019. https://www.microsoft.com/en-us/research/publication/deep-learning-methods-and-applications/?from=http://research.microsoft.com/pubs/209355/deeplearning-nowpublishing-vol7-sig-039.pdf

- Reading list for the Carnegie Mellon University course on Deep Learning, taught by Bhiksha Raj: http://deeplearning.cs.cmu.edu/ (Includes a number of classic neural network papers.)
- "11-785 Deep Learning." 11-785 Deep Learning. Accessed March 15, 2019. http://deeplearning.cs.cmu.edu/
- Maren, A. J., Pap, R. M., and Harston, C. T., 1990. Handbook of Neural Computing Applications. New York: Academic Press. Accessible via Google Books. Note: These chapters are being supplanted by new chapter drafts in Maren's Statistical Mechanics, Neural Networks, and Artificial Intelligence book.
- Reinforcement learning framework and algorithms implemented in PyTorch - By Vitchyr
- Vitchyr. "Vitchyr/rlkit." GitHub. March 01, 2019. Accessed March 15, 2019.
- "RL." Harder Choices. April 16, 2018. Accessed March 15, 2019. https://harderchoices.com/rl/