Deep and Reinforcement Learning Fundamentals

CAP5619, Fall 2018
Department of Computer Science, Florida State University

Class time and location

Monday and Wednesday, 9:30AM – 10:45AM, LOV 301 (Love Building)

Instructor

- Instructor: Xiuwen Liu
- Email: liux@cs.fsu.edu (most effective way to contact me)
- Home page: http://www.cs.fsu.edu/~liux
- Office: 168 Love Building (LOV); Phone: (850) 644-0050
- Office Hours: Monday and Wednesday, 11:00AM - 12:30PM and by appointments

Class Home Page

http://www.cs.fsu.edu/~liux/courses/deepRL/index.html.
This web site contains the up-to-date information related to this class such as news, announcements, assignments, lecture notes, and useful links to resources that are helpful. Besides the web pages, Canvas will be used to communicate changes and updates and post grades for this class; in particular, I will send emails using email addresses in the Canvas system and please make sure that your email address on record is current.

Rationale

With great advances in computer architectures, algorithms powered by high-performance computer systems are delivering superhuman performance in many areas. DeepStack, a reinforcement learning system, was able to defeat top professional poker players with a significant margin. AlphaGo was able to beat word champions in the game of Go (which has 10^{160} (billions of billions of billions of billions)^4 decision points). Face recognition systems were able to recognize faces more accurate than humans can on large databases of faces. Speech recognition systems were able to recognize speech with better accuracy than humans can. Machine translation systems can translate among multiple languages with fluent sentences. The ubiquitous connectivity enabled by the Internet has amplified their impacts quickly. Google Home and Amazon Echo devices in homes are using speech recognition techniques and car manufactures are positioned to deliver fully self-driving cars in the near future. Along with big data, these algorithms are creating the fourth industrial revolution. Almost no-cost duplication of systems and continual improvements are tipping the balances of existing approaches and fundamentally changing how we do scientific discovery and how we do our jobs. Many communities have felt the impact already. Professionals are engaging in discussions such as "lawyers could be replaced by artificial intelligence," "the end of radiology," and "machine learning: the 'next big thing' in education."

Deep learning techniques and reinforcement learning methods are the main driving force for the most successes. However, the full potential is far from being realized. For example, AlphaGo Zero has achieved an elo rating of 5,185 in the game of Go by playing against itself for 40 days. In comparison, the 1st and 100th ranked best professional players have an elo rating of 3648 and 3324 respectively (on January 19, 2018, according to https://www.goratings.org/en/), indicating the widening gap in performance. How to apply these techniques to solve many challenges the humankind is facing is a focus of research and study in many areas. This class is to cover the fundamental principles and algorithms in deep and reinforcement learning so that students will be able to develop new applications and improved algorithms that could have significant impacts, and therefore be competitive for many of jobs created by deep and reinforcement learning techniques.

Course Description

This course covers fundamental principles and techniques in deep and reinforcement learning. Topics include convolutional neural networks, recurrent and recursive neural networks, backpropagation algorithms, regularization and optimization techniques for training such networks, dynamic programming, Monte Carlo, and temporal difference, and function approximation reinforcement learning algorithms, and applications of deep and reinforcement learning. It also covers active research topics in deep and reinforcement learning areas.

Prerequisites

Senior or graduate standing in science or engineering, or permission of the instructor. Some familiarity with basic concepts in linear algebra and probability theory. Some basic knowledge of algorithm designs and some experience with Python programming.

Course Objectives

Upon successful completion of this course of study, the student will be able to:

- Implement and use backpropagation algorithms to train deep neural networks
- Apply regularization techniques to training deep neural networks
- Apply optimization techniques to training deep neural networks
- Construct and train convolutional neural networks
- Construct and train recurrent neural networks
- Implement and analyze basic deep learning algorithms for speech recognition
- Implement and analyze basic deep learning algorithms for face recognition and object recognition
- Implement and analyze basic deep learning algorithms for natural language processing
- Implement and apply policy iteration and value iteration reinforcement learning algorithms
- Implement and apply Monte Carlo reinforcement learning algorithms
- Implement and apply temporal-difference reinforcement learning algorithms
- Construct and apply on-policy reinforcement learning algorithms with function approximation

- Construct and apply off-policy reinforcement learning algorithms with function approximation

Textbook and Course Materials

Required textbook I: "Deep Learning" by Ian Goodfellow, Yoshua Bengio, and Aaron Courville (MIT Press, 2016) and it is also available online from http://www.deeplearningbook.org/; we will use it to cover deep learning. **Required textbook II: "Reinforcement Learning: An Introduction"** by Richard S. Sutton and Andrew G. Barto and we will use the second edition, which is available online from http://incompleteideas.net/book/the-book-2nd.html and should be available from MIT Press in 2018; we will use it to cover reinforcement learning.

In addition to the textbooks, I will distribute papers from the literature and notes in class along the lectures.

Student Responsibilities

In case that it is necessary to skip a class, a student is required to notify the instructor beforehand; the absence is excused if it is allowed by the University Attendance Policy (see below). The penalty for each unexcused absence is 10% reduction of attendance points (see the Grading Policy below); a student will receive 0 for attendance points if he or she has ten or more unexcused absences through the semester. In both excused and unexcused cases, the students are responsible for making up missed materials. Participation in in-class discussions and activities is also required. All submitted assignments and projects must be done by the author(s). It is a violation of the Academic Honor Code (see below) to submit other's work and the instructor of this course takes the violations very seriously.

University Attendance Policy - Excused absences include documented illness, deaths in the family and other documented crises, call to active military duty or jury duty, religious holy days, and official University activities. These absences will be accommodated in a way that does not arbitrarily penalize students who have a valid excuse. Consideration will also be given to students whose dependent children experience serious illness.

Assignments and Projects

About six homework assignments will be given along the lectures and they need to be done individually and turned in. There will be two programming projects about deep and reinforcement learning using Python and TensorFlow (or any other commonly used deep learning framework (including Keras, PyTorch, Caffe2, the Microsoft Cognitive Toolkit, Chainer, and Matlab)). There will be a midterm exam. There will also be a term project on a research topic or new algorithm implementation in deep and reinforcement learning, which is due during the final exam week.

Grading Policy

Grades will be determined as follows:

Assignment	Points	Assignment	Points
Class Attendance & Participation	10 %	Programming Project II	10 %
Homework Assignments	25 %	Midterm Exam	30 %
Programming Project I	10 %	Term Project	15 %

Grades will be computed using the weighted average as specified above and the following scale will be used (S is the weighted average on a 100-point scale):

Score	Grade	Score	Grade	Score	Grade
$93 \leq S$	A	$80 \leq S < 83$	B-	$67 \leq S < 70$	D+
$90 \leq S < 93$	A-	$77 \leq S < 80$	C+	$63 \leq S < 67$	D
$87 \leq S < 90$	B+	$73 \leq S < 77$	C	$60 \leq S < 63$	D-
$83 \leq S < 87$	B	$70 \leq S < 73$	C-	$S < 60$	F

Late Penalties

Assignments are due at the beginning of the class on the due date. Assignments turned in late, but before the beginning of the next scheduled class will be penalized by 10 %. Assignments that are more than one class period late will **NOT** be accepted.

Submission and Return Policy

All tests/assignments/projects/homework will be returned as soon as possible after grading but no later than two weeks from the due date.

Tentative Schedule

Here **Deep** refers to the "Deep Learning" textbook and **RL** refers to the "Reinforcement Learning: An Introduction" textbook.

- Week 1: Introduction (**Deep**: Chapters 1 and 5; **RL**: Chapter 1)
    - General introduction to machine learning, neural networks, deep neural networks, recurrent neural networks, and reinforcement learning
    - Successful application examples, especially in areas where superhuman performance has been achieved
    - Fundamental principles and techniques to deep learning and reinforcement learning
    - Machine Learning Fundamentals
- Week 2: Neural networks and deep feedforward neural networks (**Deep:** Chapter 6)
- Week 3: Regularization techniques for deep learning (**Deep**: Chapter 7)
- Week 4: Optimization techniques for training deep neural networks (**Deep**: Chapter 8)

- Week 5: Convolutional neural networks (**Deep**: Chapter 9)
- Week 6: Recurrent and recursive neural networks (**Deep**: Chapter 10)
- Week 7: Deep learning applications with a focus on the ones that have achieved superhuman performance (in face recognition, object recognition, speech recognition, natural language processing (machine translation)) (**Deep**: Chapter 12)
- Week 8: Reinforcement learning framework (**RL**: Chapters 2 and 3)
- Week 9: Dynamic programming algorithms for reinforcement learning (**RL**: Chapter 4)
- Week 10: Monte Carlo methods for reinforcement learning (**RL**: Chapter 5)
- Week 11: Temporal-difference learning and n-step bootstrapping algorithms for reinforcement learning (**RL**: Chapters 6 and 7)
- Week 12: Midterm Exam (Review and Exam)
- Week 13: Function approximation algorithms for reinforcement learning (**RL:** Chapters 9, 10, 11, 12, and 13)
- Week 14: Case studies of reinforcement learning applications that have achieved superhuman performance (**RL**: Chapter 16)
- Week 15: Active research topics in deep and reinforcement learning (**Deep**: Chapters 15 and 20; **RL**: Chapter 17)
- Final exam week: Term project

Academic Honor Policy

The Florida State University Academic Honor Policy outlines the University's expectations for the integrity of students' academic work, the procedures for resolving alleged violations of those expectations, and the rights and responsibilities of students and faculty members throughout the process. Students are responsible for reading the Academic Honor Policy and for living up to their pledge to "...be honest and truthful and...[to] strive for personal and institutional integrity at Florida State University." (Florida State University Academic Honor Policy, found at http://fda.fsu.edu/academic-resources/academic-integrity-and-grievances/academic-honor-policy.)

Examples of violations of the Academic Honor Policy are given in its document (found at http://fda.fsu.edu/content/download/21140/136629/AHPFinal2014.pdf). For this course, assignments/projects/exams are to be done individually, unless specified otherwise. It is a violation of the Academic Honor Policy to take credit for the work done by other people. It is also a violation to assist another person in violating the Academic Honor Policy. Concrete examples of violations include:

- ❖ Discuss the solution for a homework question
- ❖ Copy programs (done by another student) for a programming assignment
- ❖ Use and submit existing programs and reports on the world wide web for a written assignment
- ❖ Submit programs/reports/assignments done by a third party, including hired and contracted

The judgment of an alleged violation of the Academic Honor Policy will be done in accordance with the specified procedures. If the student is found to be responsible for the violation, the sanctions specified in the University Academic Honor Policy will be applied, including a zero

for the particular assignment /exam and a reduction of one letter grade in the final grade for each occurrence. In addition, administrative actions will be taken in accordance with the University Academic Honor Policy.

Americans With Disabilities Act

Students with disabilities needing academic accommodation should: (1) register with and provide documentation to the Student Disability Resource Center; and (2) bring a letter to the instructor indicating the need for accommodation and what type. Please note that instructors are not allowed to provide classroom accommodation to a student until appropriate verification from the Student Disability Resource Center has been provided. This syllabus and other class materials are available in alternative format upon request.
For more information about services available to FSU students with disabilities, contact the:

Student Disability Resource Center
874 Traditions Way
108 Student Services Building
Florida State University
Tallahassee, FL 32306-4167
(850) 644-9566 (voice)
(850) 644-8504 (TDD)
sdrc@admin.fsu.edu
http://www.disabilitycenter.fsu.edu/

Additional Information

Free Tutoring from FSU - On-campus tutoring and writing assistance is available for many courses at Florida State University. For more information, visit the Academic Center for Excellence (ACE) Tutoring Services' comprehensive list of on-campus tutoring options at http://ace.fsu.edu/tutoring or contact tutor@fsu.edu. High-quality tutoring is available by appointment and on a walk-in basis. These services are offered by tutors trained to encourage the highest level of individual academic success while upholding personal academic integrity.

Syllabus Change Policy: Except for changes that substantially affect implementation of the evaluation (grading) statement, this syllabus is a guide for the course and is subject to change with advance notice.