Learning in Humans and Machines
Cognitive Science 185:601 session 02
Computer Science 198:598 session 01
3 credits
Fall 2023

Instructor: Qiong Zhang
Instructor Email:
Classroom location: SEC 209
Meeting Times: M/W 3:00pm to 4:20pm
Instructor Office Hour: M/W 4:20pm to 4:50pm

Course Description

This interdisciplinary graduate-level course explores the parallels between human learning and machine learning. The central link between the two is the set of shared computational problems faced by humans and machines which includes making complex decisions; predicting future events; storing and retrieving information efficiently; and generalizing knowledge to new situations. By examining such problems, we will see that

1. solutions drawn on methods developed from machine learning can help us gain insights about human cognition, and conversely,
2. knowledge about how humans solve these problems can inform the development of more intelligent machines.

The first half of the course covers the application of machine learning to explain how human cognition works. We will explore the landscape of computational models of human cognition and discuss the insights these models reveal into how people learn, remember, and make complex decisions in everyday situations. The methods discussed include neural networks, symbolic approaches, Bayesian statistics, and more. The applications discussed include perception, skill learning, memory, categorization, and decision making.

In the second half of the course we will draw parallels between human learning and machine learning. Specifically, we will explore how neuroscience and our understanding of human cognition can explain and inform advances in machine learning. We will accomplish this by examining recent advances in neural networks and reinforcement learning from a psychologist’s perspective.

Each class will start with a short lecture covering the necessary machine learning techniques and cognitive science concepts to understand the readings. Following this is a student presentation of the reading. We will end with a discussion around the reading.
Learning Objectives

By the end of the course, students will
1. understand the basics of Bayesian inference, neural networks and other computational approaches,
2. understand the basics of the key aspects of human cognition such as memory and decision making,
3. be able to characterize the relationship between computational approaches to cognition and machine learning research, and
4. be able to identify ways in which computational models can be experimentally tested as models of cognition

Textbook/Resources

Lecture slides are self-contained. There is no required textbook. There will be a number of cognitive science and computer science papers for discussion, available as PDF files through the class website.

Who should take this course

The course is designed for graduate students in cognitive science, psychology, or computer science who are interested in developing computational models of human cognition and exploring the parallels between human learning and machine learning.

Coursework Requirements

Students are expected to actively participate in class discussions and sign up for at least one paper presentation (20% of total grade).

There will be a reading assignment for every class, and you are expected to arrive in class with ideas and questions to discuss. To help you develop these ideas, you are required to write short commentaries before classes— one to two paragraphs is typical (20% of total grade). A commentary might take one or several of the following forms: describe the part of the reading that you find most interesting or surprising; mention a claim that doesn’t seem right to you; describe how the work could be usefully extended; draw a connection between the reading and something else that has been discussed previously. The one-paragraph commentary should be ended with a suggestion on what would be a good discussion question to have in class. Commentaries are graded pass/fail. If you submit and pass all commentaries, you will receive full credit for this component of the course.

Students are expected to attend all classes and take notes on the most basic and important concepts discussed in each class. There will be a number of in-class quizzes distributed randomly across the semester. They consist of short true and false questions which serve as attendance and attention check for that class.
Another component of the course is an individual/team **project** to assess the student’s ability to put together the concepts and tools they have learned in the course (50% of total grade), delivered by a mid-term report, a final report, and a final presentation. The class project will be an independent research project analyzing an experiment, testing a new cognitive/machine learning model, or analyzing an existing model. The project will be an excellent opportunity for students to be engaged in multi-disciplinary research.

**Grade Evaluation**

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commentaries (due midnight prior to each class)</td>
<td>20%</td>
</tr>
<tr>
<td>Paper presentations</td>
<td>20%</td>
</tr>
<tr>
<td>In-class quizzes/attendance</td>
<td>10%</td>
</tr>
<tr>
<td>Project mid-term report</td>
<td>20%</td>
</tr>
<tr>
<td>Project final report</td>
<td>30%</td>
</tr>
</tbody>
</table>

**Schedule of Classes and Readings**

**Week 1**

**Course Overview (Sep 6)**

Review of key concepts in human cognition; history of cognitive modeling; human intelligence and machine intelligence

**Week 2**

**Marr’s three levels of analysis (Sep 11)**


**Class project briefing (Sep 13)**

Overview of class projects and datasets

**Week 3**

**Formal theories of cognition (Sep 18)**


**Rational analysis (Sep 20)**


**Week 4**

**Rational analysis (Sep 25)**

**Probabilistic models of cognition: Concept learning (Sep 27)**
Bayesian inference with a discrete space of hypotheses

**Week 5**

**Probabilistic models of cognition: Memory (Oct 2)**
Bayesian inference with a continuous space of hypotheses

**Resource-rational analysis (Oct 4)**

**Week 6**

**Resource-rational analysis (Oct 9)**

**Active learning in humans (Oct 11)**

**Week 7**

**Mechanistic models of cognition (Oct 16)**
Human decision making

**Mechanistic models of cognition (Oct 18)**
Group cognition

**Week 8**

**Neural network models of cognition (Oct 23)**
Parallel Distributed Processing
Neural network models of cognition (Oct 25)
Complementary Learning Systems

Week 9

Human-machine comparison (Oct 30)

Human-machine comparison (Nov 1)

Week 10

Inductive bias (Nov 6)

Inductive bias (Nov 8)

Week 11

Contrastive learning (Nov 13)

Nov 15 class canceled due to conference traveling.

Week 12

Brain-inspired Replay (Nov 20)

Curiosity-driven exploration (Nov 22)
Week 13
Contextual memory (Nov 27)

Hierarchical memory (Nov 29)

Week 14
Event understanding in large language models (Dec 4)

Reasoning in large language models (Dec 6)

Week 15
Final project presentations (Dec 11, Dec 13)

Academic Integrity Policies

Rutgers University regards acts of dishonesty (e.g. plagiarism, cheating on examinations, obtaining unfair advantage, and falsification of records and official documents) as serious offenses against the values of intellectual honesty. Violations of academic integrity will be treated in accordance with university policy, and sanctions for violations may range from no credit for the assignment, to a failing course grade to (for the most severe violations) dismissal from the university. Details policies can be found here: http://academicintegrity.rutgers.edu

These principles forbid plagiarism and require that every Rutgers University student:
• properly acknowledge and cite all use of the ideas, results, or words of others
• properly acknowledge all contributors to a given piece of work
• make sure that all work submitted as his or her own in a course or other academic activity is produced without the aid of unsanctioned materials or unsanctioned collaboration
• treat all other students in an ethical manner, respecting their integrity and right to pursue their educational goals without interference. This requires that a student neither facilitate academic dishonesty by others nor obstruct their academic progress (reproduced from: http://academicintegrity.rutgers.edu/academic-integrity-at-rutgers/).

Students with Disabilities
Our community values diversity and seeks to promote meaningful access to educational opportunities for all students. If you believe that you need accommodations for a disability, please follow these procedures outlined at http://disabilityservices.rutgers.edu/request.html Since accommodations may require early planning and are not provided retroactively, please initiate this process as soon as possible.

**Rutgers CS Diversity and Inclusion Statement**
Rutgers Computer Science Department is committed to creating a consciously anti-racist, inclusive community that welcomes diversity in various dimensions (e.g., race, national origin, gender, sexuality, disability status, class, or religious beliefs). We will not tolerate micro-aggressions and discrimination that creates a hostile atmosphere in the class and/or threatens the well-being of our students. We will continuously strive to create a safe learning environment that allows for the open exchange of ideas while also ensuring equitable opportunities and respect for all of us. Our goal is to maintain an environment where students, staff, and faculty can contribute without the fear of ridicule or intolerant or offensive language. If you witness or experience racism, discrimination micro-aggressions, or other offensive behavior, you are encouraged to bring it to the attention to the undergraduate program director, the graduate program director, or the department chair. You can also report it to the Bias Incident Reporting System http://inclusion.rutgers.edu/report-bias-incident/