Department of Psychology – Columbia University

Machine Intelligence

GU4236 - Fall 2020
4 points

Instructor: Trenton Jerde, Ph.D.
Class Meets: Wednesday: 6:10 – 8:00 PM
Room: 405 Schemerhorn

Office:
Office Hours:
Email: taj2128@columbia.edu

Course bulletin description

This course will survey historical and modern developments in machine intelligence from fields such as psychology, neuroscience, and computer science, and from intellectual movements such as cybernetics, artificial intelligence, neural networks, connectionism, machine learning, and deep learning. The emphasis is on the conceptual understanding of topics. The course does not include, nor require background in, computer programming and statistics. A crucial aspect of the seminar is for students to become informed consumers of applications of artificial intelligence.

Prerequisites

Undergraduate students: Undergraduate students should have taken PSYC UN1001 The Science of Psychology or an equivalent introductory psychology course. In addition, some prior coursework or experience in statistics and research methods, as well as cognitive psychology or cognitive neuroscience would be very helpful. The instructor's permission is also required.

Graduate students: Open to Ph.D. students in the Psychology department and graduate students in related departments such as computer science, neuroscience, and engineering, with instructor's permission.

Instructor

The instructor is a senior editor at *Nature Machine Intelligence*, a scientific journal on artificial intelligence, machine learning and robotics. He has taught at the University of Minnesota (cognitive science), New York University (psychology), and Columbia University (applied analytics). He has a PhD in neuroscience from the University of Minnesota and a BA in psychology from the University of Iowa.

Detailed description of the course

The goal of the seminar is to arrive at a deeper understanding of machine intelligence.
Students will lead class discussions by preparing slide presentations focused on the required readings. Students should try to engage the class as much as possible — remember, seminars are meant to be discussions.

Specifically:

- We will ask fundamental questions, such as: What is intelligence? Can machines think? Can machine intelligence approach the human level (general artificial intelligence) or surpass it (superintelligence)?
- We will distinguish data, algorithms, compute power, and technology in applications of machine intelligence, in media reports, in films, etc.
- We will compare intelligence in different kinds of machines, such as software, simulated agents, bots, drones, and robots.
- We will explore how basic psychological processes, such as perception, learning, language, thought, emotion, and movement, are embedded in machines.
- We will examine applications of machine intelligence in society, such as facial recognition, and consider the roles of tech companies such as Facebook and Google.
- We will investigate ethical issues, such as the implications of machine intelligence on privacy, surveillance, algorithmic fairness, elections, and international affairs.
- We will discuss how machine intelligence is portrayed in the media and in popular culture, and what impact these portrayals have on how the public views machine intelligence.

Learning objectives

- This course will give you training in reading primary research articles and review papers, the majority of what scientists read. Primary research articles and review papers are a much different reading experience than textbooks, and reading, dissecting, and critically thinking about them is a key skill in the understanding of science and technology.
- This course will also enable you to engage in constructive scientific conversations and debates, which will broaden and deepen your understanding of research in machine intelligence, and teach you about questions you should be asking as you read and hear about new results, and how to interpret them in the context of other studies.
- Furthermore, this course will improve how you communicate, with both oral presentations and written work.
- You will learn how to review the literature and find relevant peer-reviewed papers, allowing you to keep up to date in any field of science. You will also gain a deep understanding of the links between different areas of machine intelligence, and gain an appreciation for the development of this research over several decades, the challenges to this research (such as technological impediments and “AI winters”), and the latest advancements.

Role in the psychology curriculum

PSYC GU42xx is an advanced seminar, designed particularly for undergraduates who are majoring in Psychology or in Neuroscience and Behavior, for students participating in the Post-bac Psychology Program, and for Psychology Graduate Students. Students with a background in the computational sciences, engineering, and philosophy are also welcome to apply.

- For students pursuing the Psychology major or the Post-bac Certificate Program in Psychology, this course can be used to fulfill the advanced seminar requirement and/or the Group I Cognition & Perception distribution requirement.
- For students pursuing the Neuroscience & Behavior major, this course can be used to fulfill the P5 Advanced Seminar in Psychology requirement.
- Graduate students seeking to use this course towards M.A. or M.Phil. requirements must obtain prior permission from the Director of Graduate Studies.

Readings: There is no textbook required for this course

Readings will comprise scientific articles, book chapters, literature reviews, and commentaries in the fields of machine intelligence. PDFs of all readings will be available through CourseWorks/Canvas.

Course requirements

1. Class preparation and participation: The assigned readings are designed to expand your knowledge on the latest advancement in the field of machine intelligence and to hone your critical thinking skills. The topics discussed during the seminars are complex, leaving plenty of space to discuss and debate. Strong preparation and participation will enable us to have high-level, thought provoking discussion.

2. The day before each class period, you will be asked to submit a short (one-paragraph) reading response to CourseWorks by 9:00pm. Goals of these reading responses are to help you keep current on course topics and to help me understand where students may have had difficulty with the readings and which topics students were most intrigued by and, therefore, which areas may warrant more focus during class time. Each reading response should be no more than a short paragraph, either discussing something interesting you found in the readings or asking substantive questions about concepts in the reading you found challenging. As the goal of these assignments is to keep you up to speed and to help guide my teaching and our class discussions, the assignments will just be graded on a pass/fail basis. (I can only accept responses submitted before the deadline.)

3. Thorough reading enables thoughtful discussion. It is important to engage with the material during class discussions, since your active participation in these discussions will contribute to your final grade. If you feel that regularly contributing to class discussions is difficult for you, you should raise this issue with me as soon as possible. In such cases, we might be able to work out a way for you to participate thoughtfully through your reading responses.

Generally speaking, effective class preparation and participation could include:

- Asking insightful or clarifying questions.
- Connecting the reading to other reading we've done in the course or reading you've done on your own, drawing parallels and/or contrasts among findings.
- Actively listening to fellow classmates and responding to their ideas.
- Offering thoughtful critiques of the research methodology and providing suggestions for how it might be improved.
- Bringing in outside sources – potentially from the news media or other sources – that shed light on machine intelligence findings or that illustrate ways in which these findings are interpreted and applied.

4. Leading discussions: You will be responsible for presenting an article and leading a class discussion for at least two class meetings. I'll provide more information of the sort of presentation I'm looking for in the first week of class. But, briefly, you'll walk us through your assigned article, describing the methods and results, highlighting any strengths or weaknesses of the paper, and giving your thoughts on the meaning and importance of it. I'll ask you to complete a handout and email that to me at least two days before the date of your presentation, so that I can provide feedback in advance of your actual presentations. As the goal is for you to become more skilled in presenting research findings and leading discussions, the second presentation will be weighted more heavily than the first in calculating grades.

5. Research paper: The culmination of this course is the creation of a novel research paper relating to the material of the class. Good writing is good thinking, and a primary goal of this assignment is to help you hone your writing and critical thinking skills. The process of writing the research paper follows three steps:

a. In the middle of the course, you will be asked to identify a topic related to the class. As soon as you identify it, you are expected to email me stating your research topic, so that together we can decide whether it is appropriate. Such topic proposals should include a short paragraph about your intended topic and a list of at least five (and no more than 10) references you intend to use. I will make suggestions regarding focus, potential sources, etc.
b. Once your topic is approved, you will begin work on a first draft of the paper. Generally, you want to choose a topic that is appropriately narrow to address in an 8-10 pages paper (not including references). The paper will first introduce the topic, then review recent knowledge and advancements in the field, and then discuss future directions / breakthroughs you identify.
c. Towards that end, I will provide comments and suggestions on your first draft, and you will be expected to make substantive changes – not just copyediting, but rather larger edits, such as reworking entire sections, drawing on new sources, and providing more analysis. The final draft of the paper will be graded not only as a standalone paper but also in how it demonstrates improvement upon the earlier draft.

Grading

Grades will be calculated based on the percentages outlined below.

1. Class preparation and participation ...……25%
    a. Reading responses 10%
    b. Contribution to class discussion 15%
2. Discussion leading ……30%
    a. First presentation 10%
    b. Second presentation 20%
3. Research paper ….……45%
    a. Proposal 5%
    b. First draft 10%
    c. Final draft 30%

Schedule and required readings

Session	Topics and Assignments	Readings
Week 1	Introduction to the seminar: review of the syllabus and overview of the topic	• Interview with Yoshua Bengio, in *Architects of Intelligence* (2018), ed. Martin Ford. • Interview with Yann LeCun, in *Architects of Intelligence* (2018), ed. Martin Ford. • Interview with Geoffrey Hinton, in *Architects of Intelligence* (2018), ed. Martin Ford.
Week 2	Foundations I	• Turing, A.M. (1948), Intelligent machinery. Chapter 10 in *The Essential*

		Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life, Plus the Secrets of Enigma (ed. J. Copeland. Clarendon, 2004. • Turing, A.M. (1950). Computing Machinery and Intelligence. *Mind* 59 (236): 433–460. • Piccinini, G. (2004). The first computational theory of mind and brain: A close look at McColloch and Pitts's "Logical Calculus of Ideas Immanent in Nervous Activity". *Synthese* 141.
Week 3	Foundations II	• Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*. • Minsky, M. (1961). Steps towards artificial intelligence. In *Computers and Thought*, ed. E.A. Feigenbaum and J. Feldman. • Marr, D. (1982). The three levels. In Chapter 1 of *Vision*.
Week 4	Neural networks I	• Mitchell, M. (2019). Neural networks and the ascent of machine learning. Chapter 2 in *Artificial Intelligence: A Guide for Thinking Humans*. • Sejnowski, T. (2018). The dawn of neural networks. In *The Deep Learning Revolution*. • Interview with Geoffrey Hinton (1998). In *Talking Nets, An Oral History of Neural Networks*, eds. J.A. Anderson and E. Rosenfeld. • Interview with Carver Mead (1998). In *Talking Nets, An Oral History of Neural Networks*, eds. J.A. Anderson and E. Rosenfeld.
Week 5	Neural networks II	• Mitchell, M. (2019). ConvNets and ImageNet. Chapter 5 in *Artificial Intelligence: A Guide for Thinking Humans*. • Rumelhart, D., Hinton, G. & Williams, R. (1986). Learning representations by back-propagating errors. *Nature* 323, 533–536. • Papert, S. (1988), One AI or many? In *The Artificial Intelligence Debate*, ed. S.R. Graubard.

Week 6	Reinforcement learning	• Mitchell, M. (2019). Rewards for robots. Chapter 8 in *Artificial Intelligence: A Guide for Thinking Humans*. • Mitchell, M. (2019). Game on. Chapter 9 in *Artificial Intelligence: A Guide for Thinking Humans*. • Sutton, R.S., Barto, A.G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychological Review* 88(2): 135-70.
Week 7	Deep learning	• Bengio, Y. (2016). Machines who learn. *Scientific American*. • LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature* 521, 436–444.
Week 8	Psychological considerations in machine intelligence	• Lake, B. et al. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*. • Commentaries by other experts on this article and the authors' replies.
Week 9	Fundamental issues: Learning vs. built-in priors; nature vs. nurture; rationalist vs. empiricist	• Sutton, R. (2019). The bitter lesson. • Papert, S. (1980). Artificial intelligence and general developmental mechanisms. Discussion by Noam Chomsky, Seymour Papert, Scott Atran. In *Language and Learning: The Debate Between Jean Piaget and Noam Chomsky*, ed. M. Piattelli-Palmarini.
Week 10	Fundamental issues: Reverse engineering the mind; artificial general intelligence	• Interview with Josh Tenenbaum, in *Architects of Intelligence* (2018), ed. Martin Ford. • Interview with Demis Hassabis, in *Architects of Intelligence* (2018), ed. Martin Ford.
Week 11	Perspectives on AI, machine learning, and deep learning I	• Gopnik, A. (2019). AIs versus four year olds. Chapter 21 in *Possible Minds: 25 Ways of Looking at AI*, ed. John Brockman. • Marcus, G., Davis, E. (2019). Minding the gap. Chapter 1 in *Rebooting AI: Building Artificial Intelligence We Can Trust*. • Marcus, G., Davis, E. (2019). Deep learning, and beyond. Chapter 3 in *Rebooting AI: Building Artificial Intelligence We Can Trust*.
Week 12	Perspectives on AI, machine learning, and deep learning II	• Hofstadter, D. (2018), The shallowness of Google Translate. *The Atlantic*

		• Katz, Y. (2012), Noam Chomsky on where artificial intelligence went wrong. *The Atlantic* • Jordan, M.I. (2018), Artificial intelligence – the revolution hasn't happened yet. *Medium*.
Week 13	Ethical issues: face recognition, algorithmic fairness, rise of the tech companies	• Wee. S-L., Mozur, P. (2019), China uses DNA to map faces, with help from the West. *New York Times* • Lee, N.T., Resnick, P., Barton, G. (2019), Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. Brookings Report.
Week 14	Machine intelligence in the media and pop culture; students summarize papers	• We will discuss Black Mirror, The Matrix films, 2001: A Space Odyssey, Her, The Terminator; and writings on AI, including on superintelligence.

Class policies: Important information below – please read carefully

Special needs: If you are a student with special needs and require any type of accommodation, please make an appointment with me before the first class to discuss your needs. You should also contact the office of Disability Services (https://health.columbia.edu/disability-services) before the first class to register for specific accommodations. If you have problems reading specific kinds of text (e.g., based on font or text size), please see me so I can make you exams (and a syllabus,
and anything else you need) that you can more easily read.

Religious observances: If you are going to miss class(es) due to religious holidays, you must notify me during the first week of class so that accommodations may be made.

Academic integrity: As members of this academic community, we are responsible for maintaining the highest level of personal and academic integrity: "Each one of us bears the responsibility to participate in scholarly discourse and research in a manner characterized by intellectual honesty and scholarly integrity.… The exchange of ideas relies upon a mutual trust that sources, opinions, facts, and insights will be properly noted and carefully credited. In practical terms, this means that, as students, you must be responsible for the full citations of others' ideas in all of your research papers and projects… [and] you must always submit your own work and not that of another student, scholar, or internet agent" (from the Columbia University Faculty Statement on
Academic Integrity) http://www.college.columbia.edu/academics/academicintegrity . Cheating and plagiarism – whether intentional or inadvertent – is a serious violation of academic integrity. Plagiarism is the practice of claiming or implying original authorship of (or incorporating materials from) someone else's written or creative work, in whole or in part, without adequate acknowledgement. If you have any questions about what constitutes plagiarism and/or how to properly cite sources, please come to me. I am more than happy to help. Similarly, if you put yourself in a situation in which you think your best option might be to cut some corners, see me. If you feel like you are falling behind, don't understand the material, or are not confident about your ability to take tests, talk to me as soon as possible instead of taking measures that go against

principles of academic integrity. We are here to learn, not to merely judge. It is a far better option to come talk to me than compromise your academic integrity and potentially put your academic standing in jeopardy.

Sexual Respect: Any form of gender-based misconduct will not be tolerated. Columbia University is committed to fostering an environment that is free from gender-based discrimination and harassment, including sexual assault and all other forms of gender-based misconduct. Visit this website for more information: http://sexualrespect.columbia.edu/

Attendance: Coming to class is meaningless if class time is spent inappropriately. Chatting with friends, watching videos online, and browsing social media are not appropriate activities for the classroom. Also, remember to silence your cell phone before class. Generally, eliminate distractions as much as possible to respect your classmates, as well as increase your chance of staying focused and learning the material during class.

Writing Center: I encourage you to visit the Writing Center, where you can receive free individual consultations on your writing at any stage in the writing process, including brainstorming. Writing consultants work with all members of the Columbia community on any academic or nonacademic writing. You can make an appointment and view drop in hours on their website [www.college.columbia.edu/core/uwp/writing-center].

Changes to the Syllabus might happen during the course. The most recent version will always be posted to CourseWorks.