GenEd 1125: Artificial and Natural Intelligence - Syllabus

What does it mean for a machine to be intelligent, how does current artificial intelligence compare with animal intelligence, and should we be worried?

SCIENCE & TECHNOLOGY IN SOCIETY

Instructor: Professor Venkatesh Murthy
Raymond Leo Erikson Life Sciences Professor of Molecular and Cellular Biology
Director of the Center for Brain Science

What is intelligence? An inquiry into the nature of intelligence can take different forms – philosophical, biological, mathematical or technological. In this course, we will use machine intelligence (everything from voice recognizing smartphones to game-playing computers) as a handle to think about natural intelligence (brains and behavior of animals). Although we will start with big, general questions, we will quickly move to concrete queries about brains and computers. This approach, rather than just starting with brains of animals, may be useful in framing more universal questions independent of the specific architecture of brains of animals. As machines increasingly perform tasks that were once thought to be solely in the domain of humans, there is an urgent need for discussions of the moral and societal implications of artificial intelligence. This course targets students interested in brains and computers in equal measure, and who are comfortable discussing ethical concerns.

Teaching Fellows (TFs)

- Kumaresh Krishnan (Head TF) | kumaresh_krishnan@g.harvard.edu
- Alex Chen | abchen@g.harvard.edu
- Hannah McCalmon | hwatson@g.harvard.edu
- Leannah Newman | leannah_newman@g.harvard.edu
- Marc Duque Ramirez | mduqueramirez@g.harvard.edu
- Nico Herndon | nherndon@g.harvard.edu
- Stephan Foianini | s_foianini@fas.harvard.edu

Course Prerequisites

There are no formal prerequisites to this course, and students do not need to be from a technical or science background. Some basic high school math may be required for some problem sets. Problem sets will be a mix of short-form discussion questions and simple technical problems.

Assessment

10% - Lecture/Section attendance and participation*
10% - Readings

40% - Problem sets (4 total), 10% each
20% - Mid-term in class exam
20% - Final paper (2+3+15 across deliverables)

*Lectures will be interactive, with discussions throughout. Lectures will also be live-recorded and released after each class.

Summary of Important Dates

All problem sets due on Friday at 11:59pm

- 01/31 - PSet 1 handed out (Lectures 1-4)
- **02/9 - PSet 1 due @ 11:59pm**
- 02/14 - PSet 2 handed out (Lectures 5-8)
- 02/20 - Deadline to add / drop a course (fee) and to change from letter-graded to Pass/Fail
- **02/23 - PSet 2 due @ 11:59pm**
- 03/02 - Midterm (take-home) handed out
- **03/06 - Midterm (in class)**
- 03/09 to 03/17 - Spring Break!
- 03/27 - PSet 3 handed out (Lectures 13-16)
- **04/05 - PSet 3 due @ 11:59pm**
- 04/10 - PSet 4 handed out (Lectures 17-20)
- **04/19 - PSet 4 due @ 11:59pm**
- **05/06 - Final paper due**

Late policy: You are allowed one (1) late pass for problem sets (up to two days). After that, 1 point will be deducted off your problem set grade for each subsequent day that it is late.

Collaboration and Academic Integrity

Discussion and the exchange of ideas are essential to doing academic work. For assignments in this course, you are encouraged to consult with your classmates as you work on problem sets and exercises. However, after discussions with peers and/or TFs, make sure that you can work through the problem yourself and ensure that any answers you submit for evaluation are the result of your own efforts. If any books, articles, websites, lectures, etc that have significantly helped you with your work, please use appropriate citation practices. Be sure to familiarize yourself with the Harvard honor code, and follow it.

Academic Accommodations

Any student needing academic adjustments or accommodations is requested to present their letter from the Disability Access Office (DAO) and speak with the professor by the end of the second week of the term. Failure to do so may result in the course head's inability to respond in a timely manner. All discussions will remain confidential, although DEO may be consulted to discuss appropriate implementation.

Course Schedule

Week 1: Introduction

Monday, 01/22 — Lecture 1: What is Intelligence?

- Course introduction and overview
- What do we consider intelligent behavior? Can intelligence be reduced to a computation? Does it depend on the hardware performing the computation? What are the differences and commonalities between

artificial learning systems and biological brains? We will use the example of visual processing to introduce Marr's levels of understanding: Computation, Algorithm, and Implementation.

Wednesday, 01/24 — Lecture 2: With Great Power Comes Great Responsibility

- Over the past decade, our understanding and ability to measure brain signals has progressed tremendously. Similarly, artificial learning systems are now ubiquitous and can perform tasks previously thought to be reserved to humans. How do we use this knowledge? We will begin to discuss the moral, societal, and ethical implications of these advances and introduce the topics we will cover later in the course.

Week 2: The Building Blocks of Biological and Artificial Intelligence

Monday, 01/29 — Lecture 3: The Building Blocks of the Brain

- We will cover the basics of neurophysiology (single neuron, action potential and synapses). We will begin to discuss how neurons are connected to each other to perform certain functions – simple examples of pattern generators. We will discuss the similarities and differences between neural and AI architecture.

Wednesday, 01/31 — Lecture 4: Guest Lecturer

- What are brains made of and how can we look into them?
- Problem Set 1 handed out (Lectures 1-4)

Week 3: The Building Blocks of Biological and Artificial Intelligence (cont'd)

Monday, 02/05 — Lecture 5: The Architecture of the Brain - I

- We will discuss how neurons form brain circuits and brain areas with specific functional roles. We will describe how we can measure the computations performed by the brain. We will discuss new technologies and how they in turn depend on advances in machine learning. We discuss how the concept of modularity is useful in AI.

Wednesday, 02/07 — Lecture 6: The Architecture of the Brain – II

- We extend the foundation from the previous lecture to discuss further neuroscience centered aspects on neurons, connections, information flow and networks.

Friday, 02/09 — Problem Set 1 due @ 11:59pm

Week 4: The Building Blocks of Biological and Artificial Intelligence (cont'd)

Monday 02/12 – Lecture 7: The Building Blocks of Neural Networks

- We will introduce the perceptron model and the concept of linear separability. We will discuss the limitations of the perceptron and the problem of credit assignment in large networks. [Demonstration: Perceptrons and simple classifiers]

Wednesday, 02/14 — Lecture 8: How Networks Learn

- We will introduce the backpropagation algorithm which is one the essential methods that allow neural networks to learn. We will also discuss the importance of datasets in learning large models that perform complex tasks. [Demonstration: digit recognition, MNIST].

- Problem Set 2 handed out (Lectures 5-8)

Week 5: Learning to Represent

Monday, 02/19 — President's Day

- No class!

Wednesday, 02/21 - Lecture 9: Dimensionality Reduction and Invariant Representations

- We will introduce the notions of invariant representations and dimensionality reduction. Using the example of face representations, we will look at neural representations in inferior temporal cortex and in artificial neural networks. [Demonstration: image recognition in CNNs]

Friday, 02/23 — Problem Set 2 due @ 11:59pm

Week 6: Learning Complex Representations

Wednesday, 02/21 — Lecture 10: Primary Visual Cortex and Convolutional Neural Networks

- Starting with the example of visual recognition, we will introduce the neurobiology of early visual processing (retina/V1). We will then introduce convolutional neural network (CNN) architectures and examples of their success.

Wednesday, 02/28 — Lecture 11: Guest Lecturer

Week 7: Representation To Action

Monday, 03/04— Lecture 12: Moving In The World

- How do we use the networks and techniques in earlier lectures to complete tasks in the real world?

Wednesday, 03/06 — Mid-term Exam (In class)

- Open notes
- Tests more application and very minimal fact recall

Spring Recess
03/09 to 03/17 — Spring Recess

Week 8: Learning In The Brain
Monday, 03/21— Lecture 13: Types Of Learning

- What does it mean to learn and what are the different ways by which this is achieved? How does the brain accomplish associative learning?

Wednesday, 03/20 — Lecture 14: Guest Lecturer

Week 9: Language And Learning
Monday, 03/25— Lecture 15: Natural Language

- What forms of communication do animals use with each other? We will look at examples across the animal kingdom: ants, songbirds, mouse vocalization, human speech. Is human language qualitatively different? What are the human brain areas dedicated to language? Are they unique to humans?

Wednesday, 03/27 — Lecture 16: Language and Machines

- What does understanding language mean? Is it simply manipulating symbols or is there something else? Recent language models (google translate, GPT-3) work well in real world applications but they make unnatural mistakes. Is "something" still missing?
- Problem Set 3 handed out (Lectures 13-16)

Friday – 03/29 – Final paper deliverable 1 due

Week 10: Applications

Monday, 04/01— Lecture 17: Learning to Play Complex Games

- Tic-tac-toe, Chess, Go, video games: humans have long designed games requiring planning. Artificial systems can now beat humans in all of these. We will ask what is behind these algorithms. Can we solve them by looking at all possible moves? How much is built in and how much is learned?

Wednesday, 04/03 — Lecture 18: Learning How To Drive A Car

- Automated driving has long been a goal of artificial intelligence. We will give examples of early systems and ask what the components needed to perform in a "natural" environment" are.

Friday, 04/05 – Problem Set 3 due @ 11:59 pm

Week 11: Understanding & Leveraging The Black Box?

Monday, 04/08— Lecture 19: Mind-Reading

- Recent advances in non-invasive neurotechnology allow us to decode brain activity with increasing precision. This offers exciting clinical opportunities through neural prostheses to recover lost abilities such as speech or walking. But commercial applications also raise questions about the ownership and privacy of thoughts. How do we interpret from machines and artificial neural networks?

Wednesday, 04/10 — Lecture 20: Transfer Learning and Generalizability

- Do we build targeted algorithms for each task from scratch? How can we transfer the knowledge from one task to another, just as humans (and some other animals) often do?
- How do we conceptualize moving towards a generalizable model and is this a reasonable ask of neural networks? What does it take to 'replicate the brain'?
- Problem Set 4 handed out (Lectures 17-20)

Week 12: Biases and Reality

Monday, 04/15— Lecture 21: Human and Algorithmic Biases

- Human (animal) perception and decision making is rife with illusions and biases – some examples are size illusions, implicit biases etc. Why and how do these arise? How do priors (assumptions) influence these? How do we reconcile these lapses and biases with "intelligent" behavior?
- **Final paper Deliverable 2 due @ 11:59 pm**

Wednesday, 04/17 — Lecture 22: Generative AI

- Content is produced at large scales in the current day. How to distinguish AI generated and naturally sourced content? How do we differentiate between real and fake? How blurred are the lines and what are the implication on ethics and policy?
- What are popular techniques in the space of Generative AI?

Friday, 04/19 – Problem Set 4 due @ 11:59 pm

Week 13: Final Thoughts

Monday, 04/22— Lecture 23: Guest Lecturer

Wednesday, 04/24 — Course Summary & Feedback

Friday, 05/06 — Final paper due @ 11:59pm