Introduction to Artificial Intelligence

Lecture 1: Introduction / Course content

September 4, 2025

Course introduction

- **This semester: An introductory class on artificial intelligence**

- **Intelligence:** For thousands of years, we have tried to understand *how we think and act*. How does our brain perceive, understand, predict, and survive in a world larger and more complicated then the brain?

- **Artificial intelligence (AI):** Not just understanding but also *building* intelligent entities---machines that can compute how to act effectively and safely in new environments

Why learning about AI?

- **Economy:** In 2024, U.S. private AI investment grew to $109.1 billion
  - 12 times China's $9.3 billion and 24 times the U.K.'s $4.5 billion
  - AI business usage is also accelerating: 78% of organizations reported using AI in 2024, up from 55% the year before
  - Globally, corporate AI investment reached $252.3 billion in 2024, with private investment climbing 44.5% and mergers and acquisitions up 12.1% from the previous year
  - Generative AI usage in at least one business function more than doubled—from 33% in 2023 to 71% last year

- **Reference:** https://hai.stanford.edu/assets/files/hai_ai_index_report_2025.pdf

Why learning about AI?

- **Science and technology:** AI's growing importance is reflected in major scientific awards. Two Nobel Prizes recognized work that led to deep learning (physics) and to its application to protein folding (chemistry), while the Turing Award honored groundbreaking contributions to RL
  - **Complex reasoning remains a challenge**: AI models excel at tasks like IMO but still struggle with complex reasoning. They often fail to reliably solve logic tasks even when provably correct solutions exist, limiting their effectiveness in high-stakes settings where precision is critical

- **The intellectual frontier of AI are wide open:** Unlike other fields such as physics, AI still has many openings for full-time masterminds

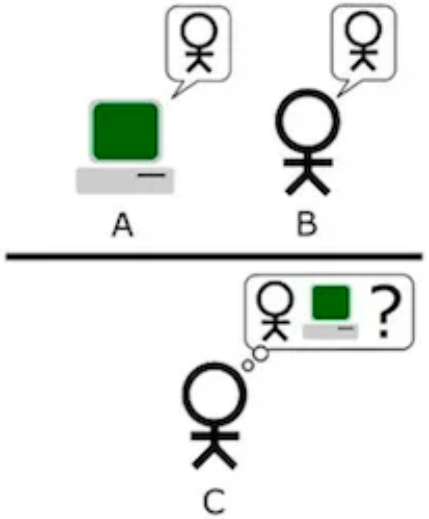
What is AI?

- Desiderata:
  - Intelligence in terms of fidelity to human performance (also known as "rationality")
  - Capable of internal thought processes and reasoning

Human	Thought
Behavior	Rational

What is AI? The Turing test approach

- **"Can a machine think?"** If a computer can fool a human into thinking he or she is conversing with another human rather than a machine, then the computer can be said to be a true example of artificial intelligence



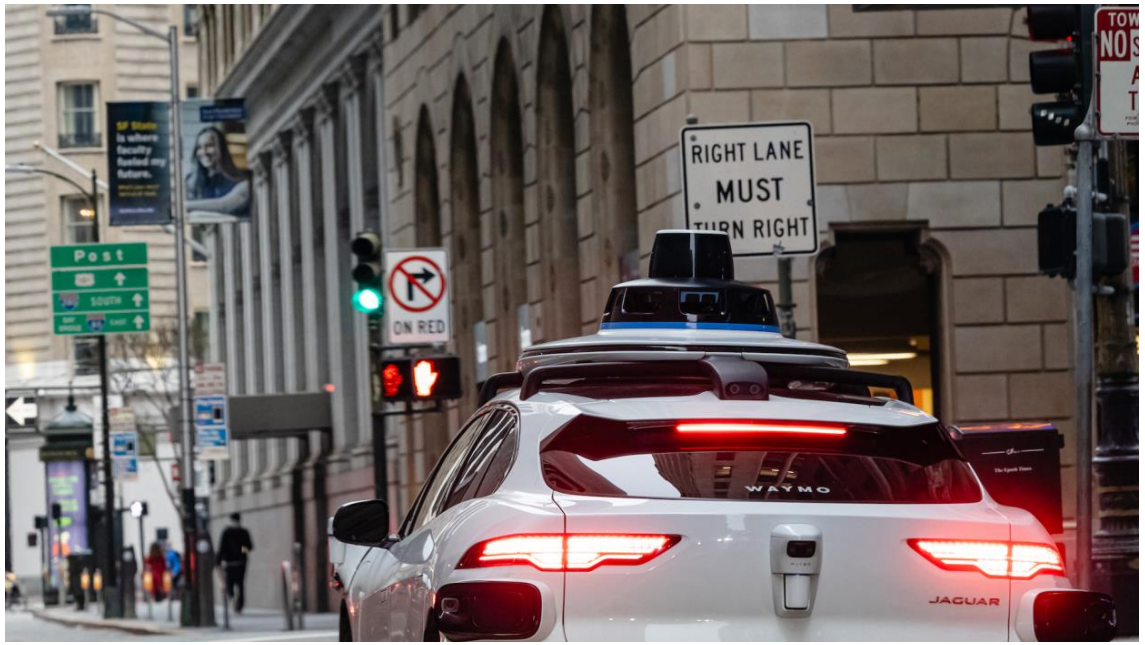
- **Required skills**: natural language processing, knowledge representation, automated reasoning, and machine learning

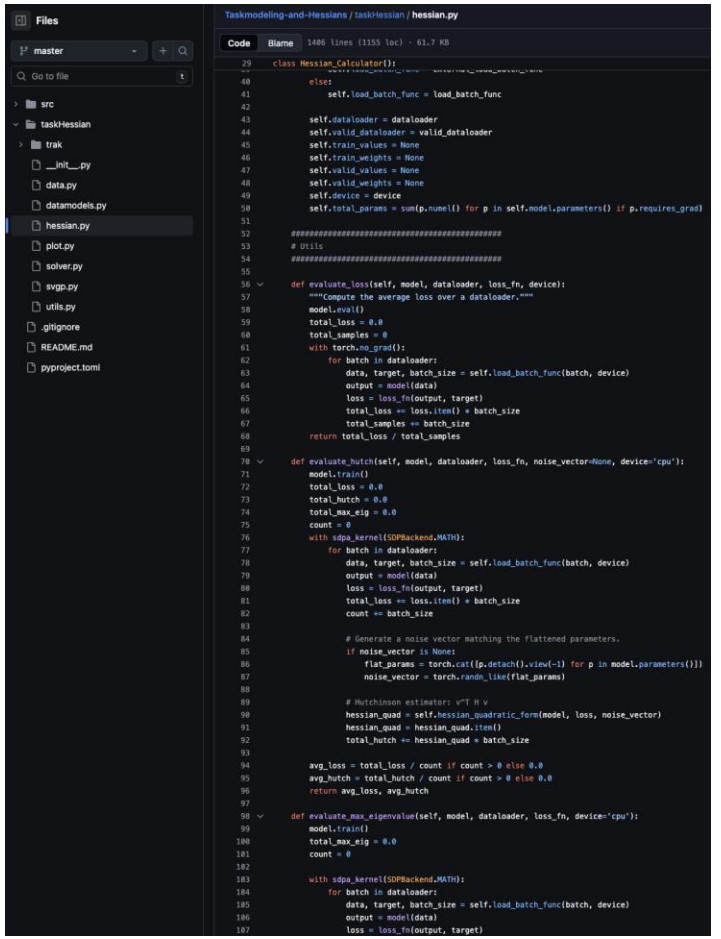
The rational agent approach

- **Agents:** operate autonomously, perceive their environment, persist over a prolonged time period, adapt to a change, create and pursue goals
  - **Self-driving cars**: sense complex scenes in real time, and make safe driving decisions
  - **Virtual LLM assistant**: understand rich subtleties of natural language, infer use intent, and generate responses

Complex real-world problems

- **Bridging the gap**





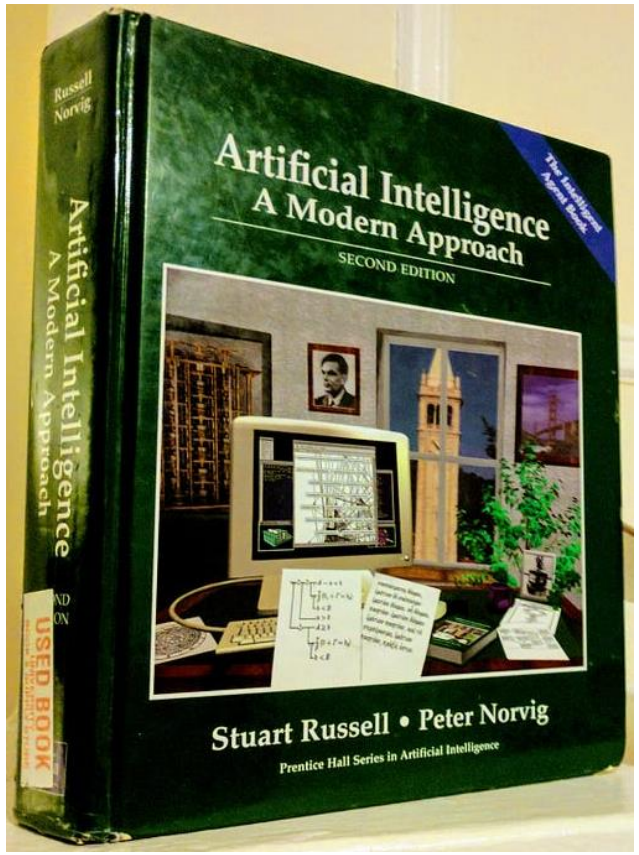
Syllabus

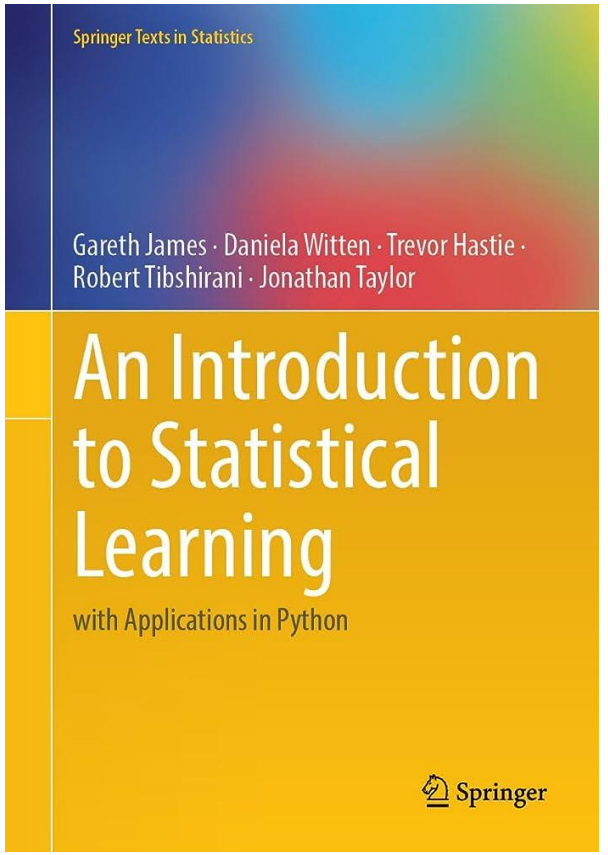
- **Part I:** Machine learning, neural networks, and language modeling
- **Part II:** Intelligent search, games, and reinforcement learning,
- **Part III**: Reasoning, knowledge representation, and uncertainty

- **Prerequisites**: Linear Algebra, Applied Probability, Python
  - **Note:** These are soft requirements in the sense I will use algebra and python at various places. However, you are not required to have taken the corresponding classes beforehand. The TAs and I will make sure to cover up the ground so you have all you need to complete this course

- **Complete syllabus, schedule, and course homepage:** see canvas

Recommended (optional) textbooks





+ Handwritten notes (from me and TAs)

Course information

- **Instructor**: Ryan Zhang
- **Location**: WVG 108
- **Time**: Mondays and Thursdays from 11:45 AM to 1:25 PM

Coursework and grading policy

- Homework 40%
  - 5 - 6 problem sets with a mix of mathematical and coding questions
- Exam 40%
  - Nov 20 – Nov 23, take home, pick 24 hours
- Course project proposal presentation + final project presentation 20%
  - In class, proposal presentation on Oct 20 and Oct 23; final project presentation on Dec 1
  - We will provide a template project around AI agents such as LLMs
  - Another option is to pick a problem you are interested and develop a machine-learning solution to solve it

Lecture plan

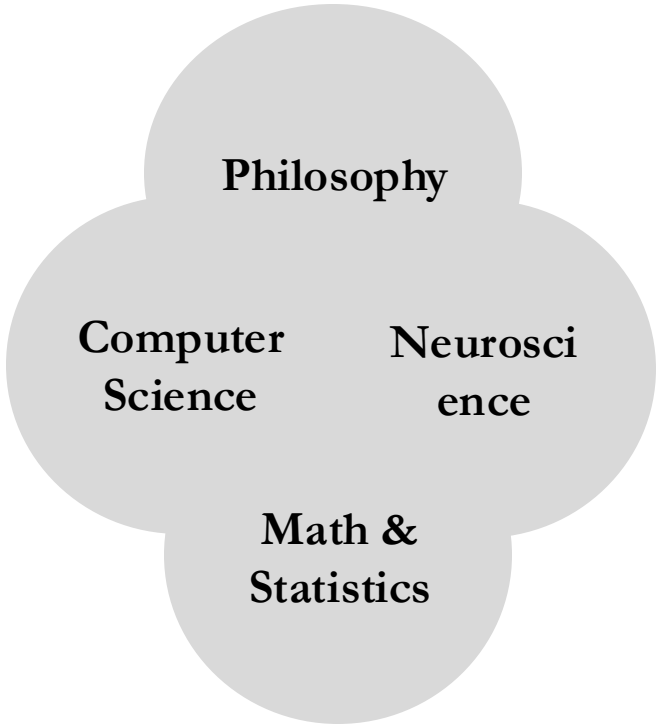
- **Several examples of machine learning**
- History of AI

Foundations of artificial intelligence

- Back to our question: how to build machines that exhibit intelligent (human-like) behavior?

- Some low-level questions:
  - What are the formal rules to draw valid conclusions?
  - What can be computed?
  - How do we reason with uncertain information?

- Some high-level questions:
  - How do brains process information?
  - How does knowledge lead to action?



Machine learning

- One of the major technical developments that helped fuel the growth of AI is machine learning and deep learning, including large language models, generative models

1st example: Rental cost prediction

- **Example**: Predict the apartment rental cost in Boston

REDFIN™

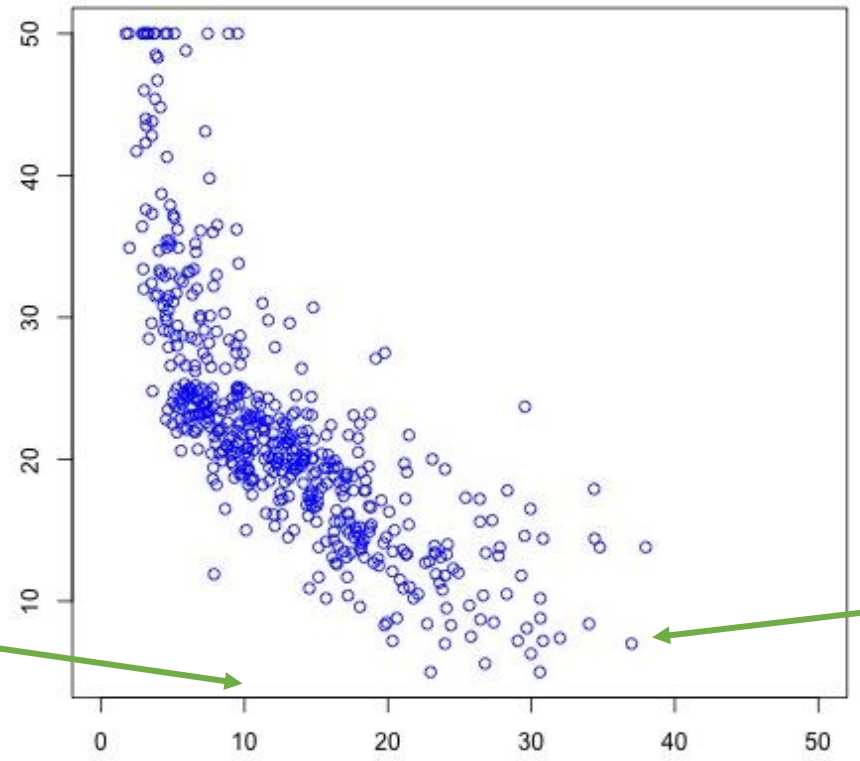
Zillow®

- **Mathematical setup:** $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(n)}, y^{(n)})$
  - $x^{(i)}$: number of bedrooms, number of bathrooms, number of square feet, zip code, distance to nearest train station, number of groceries nearby…
  - $y^{(i)}$: rental price of the apartment
- **Linear regression:** set up a linear model θ, minimize the mean squared error

$$\hat{L}(\theta) = \frac{1}{n}\sum_{i=1}^{n}\left(\theta^{\top}x^{(i)} - y^{(i)}\right)^2$$

Illustration

- This is a scatter plot of the dataset: LSTAT (income status among population vs. MEDV (median rental value)



$x = 10$
Predict y =?

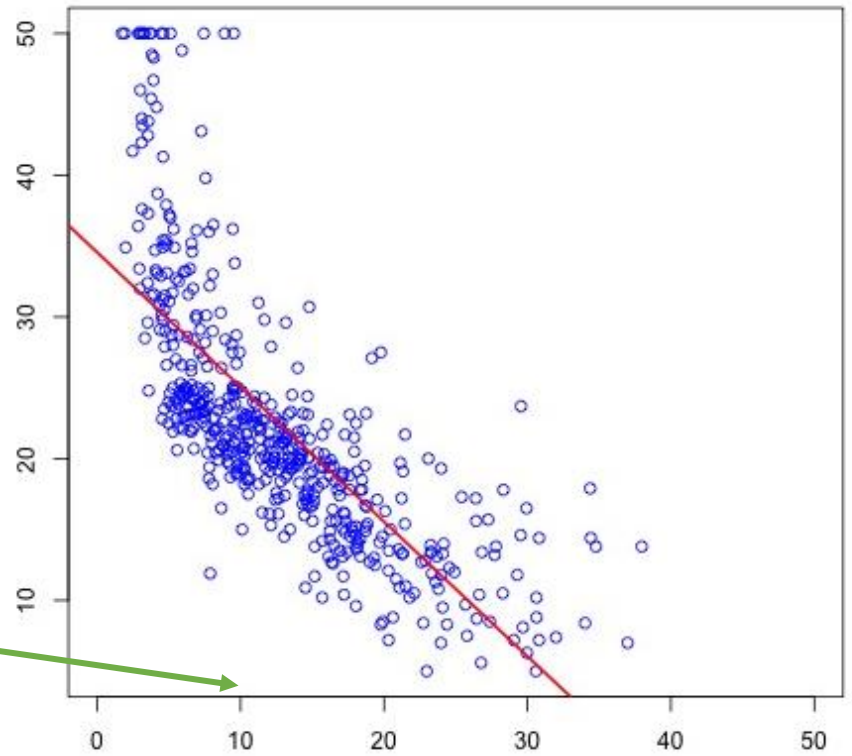
i-th sample
$(x^{(i)}, y^{(i)})$

Illustration

- **Fit a line by minimizing the MSE metric**

$x = 10$

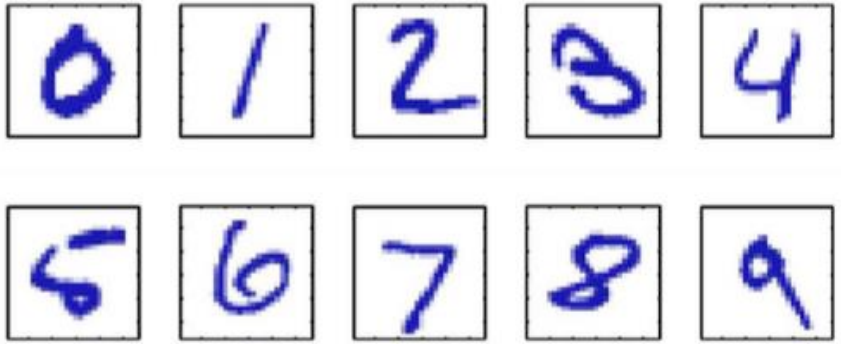
Predict $y = ?$



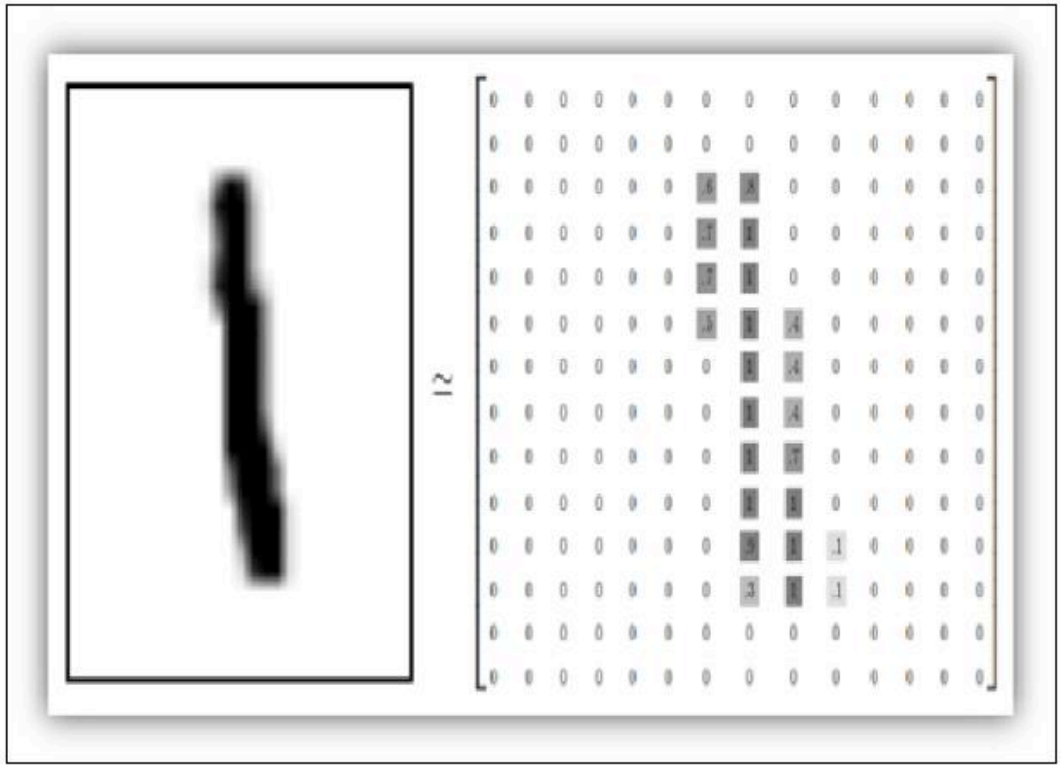
Fit a line to the data

2nd example: Classifying handwritten digits

- Each image is 28 by 28 pixels, corresponding to a 784-dimensional vector

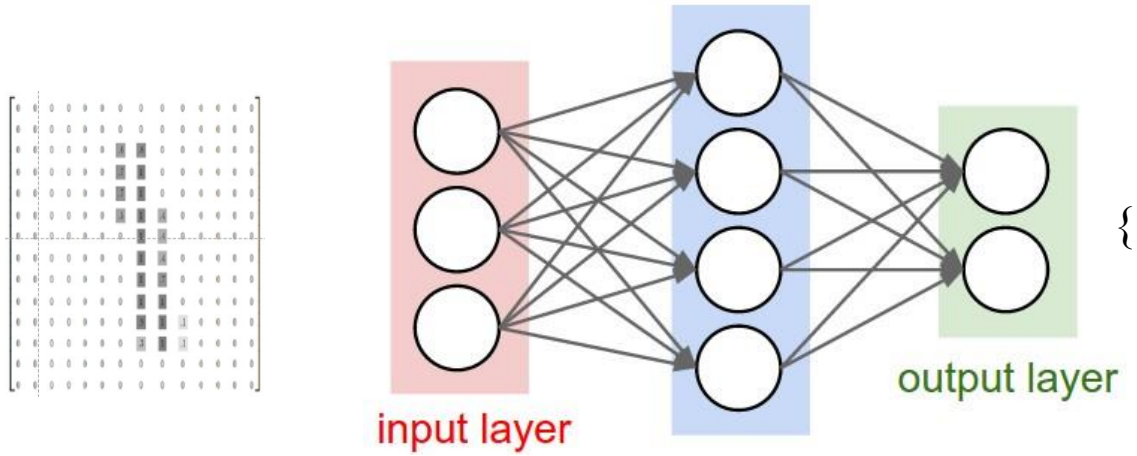


- Preprocessing: First need to map the image to an encoding: Here we'll use a matrix to represent of a black-n-white digit



Feedforward neural networks

- Let's consider the simplest form of neural network models



{0, **1**, 2, 3, 4, 5, 6, 7, 8, 9}

- **Input layer:** Receives the encoding of the handwritten digit 1
- **Hidden layer:** Maps the input encoding to a representation through a nonlinear transformation
- **Output layer:** Maps the representations to class probabilities

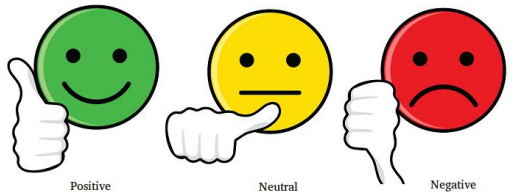
Generalization

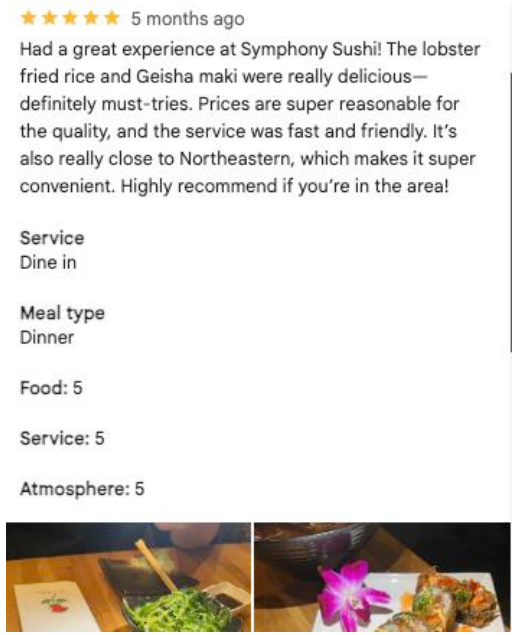
- The main conceptual difficulty of learning is that if done properly, the trained model will be able to produce good predictions beyond the set of training examples

- This leap of faith is called **generalization**, either explicitly or implicitly, and is at the heart of any machine learning algorithm
  - This can be formalized using tools from probability and statistical learning theory

3rd example: Sentiment analysis

- Given a list of words: $\{A_1, A_2, \ldots, A_n\}$, predict sentiment y which is either positive, neural, or negative





- Choose the most likely hypothesis given the list of words

$$\arg\max_{y} \Pr(A_1, A_2, \ldots, A_n | y) = \frac{\Pr(A_1, A_2, \ldots, A_n, y)}{\Pr(y)}$$

- Naïve Bayes assumes conditional independence

$$\Pr(A_1, A_2, \ldots, A_n | y) \approx \Pr(A_1|y) \cdot \Pr(A_2|y) \cdot \cdots \cdot \Pr(A_n|y) = \frac{\Pr(A_1, y)}{\Pr(y)} \cdot \cdots \cdot \frac{Pr(A_1, y)}{Pr(y)}$$

Course plan

Intelligent search
Markov decision processes

Reasoning
Constraint satisfaction
Planning
Uncertainty and Bayesian methods

Low level

Machine learning

High level

Representation learning

Natural language processing

Reinforcement learning

Virtual assistants

Robotics

Mid break

- Complete survey!! We'll synthesize the survey results and share in the next lecture

https://forms.gle/b4264df8TajGTfnw5

Lecture plan

- Several examples of machine learning
- **History of AI**

History of AI

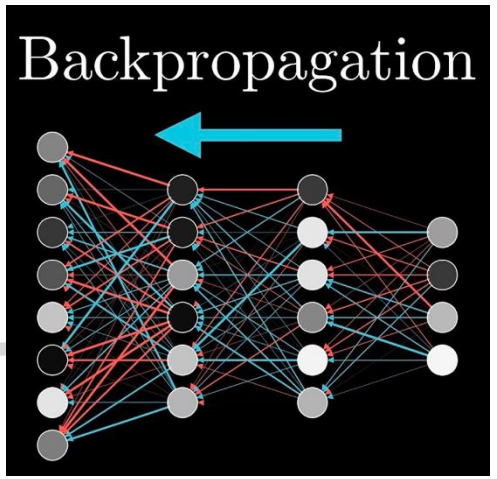
- The inception of artificial intelligence (1943 – 1956)
  - Alan Turing, "Computing Machinery and Intelligence"
  - The imitation game
- John McCarthy (Dartmouth)
  - 1956, summer workshop that includes legendary thinkers including Allen Newell, Herbert Simon, Arthur Samuel
- Newell and Simon introduced a mathematical theorem-proving system called Logic Theorist (ST), capable of proving most of the theorems in Chapter 2 of Russell and Whitehead's *Principia Mathematica*

History of AI

- Early enthusiasm, great expectations (1952 – 1969)
  - *Machines will be capable, within twenty years, of doing any work a man can do.* — Herbert Simon
  - *Within 10 years the problems of artificial intelligence will be substantially solved.* —Marvin Minsky
  - *I visualize a time when we will be to robots what dogs are to humans, and I'm rooting for the machines.* —Claude Shannon
- Then come a dose of reality (1966 – 1973)
  - Problems: **limited computation** (search space outpace hardware), **limited information** (complexity of AI)

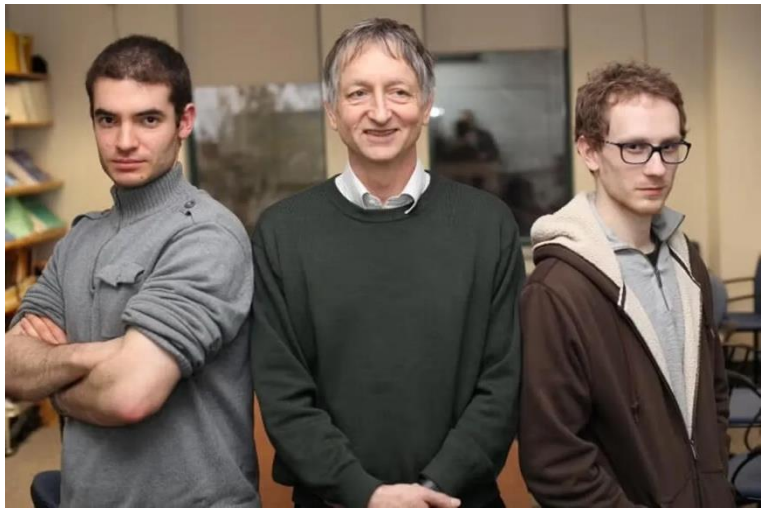
History of AI

- Expert systems, the return of neural networks, probabilistic reasoning, and machine learning
  - 1969-1986: Expert systems elicit specific domain knowledge from experts in form of rules
  - 1986-present: Neural networks become the workhorse of modern machine learning, powered by the pack-propagation learning algorithm
    - David McAllester (1998): AI and computer science can benefit from each other
    - Judea Pearl (1988): Introduce Bayesian networks to tackle probabilistic reasoning



Deep learning (2011 – present)

- 2012: AlexNet obtains huge gains in computer vision and object recognition; transformed computer vision
  - **Geoffrey Hinton** (U-Toronto) and his students Ilya Sutskever and Alex Krizhevsky



- 2016: AlphaGo uses deep reinforcement learning, defeat world champion Lee Sedol in the Go game

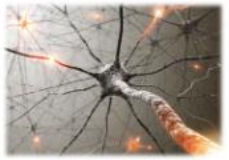
GPT models

- 2023: **General pretrained transformer** (GPT) models such as GPT-3, Claude, Gemini begin to impact the world

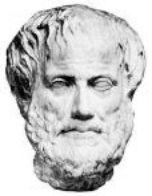
- **Multitasking**: Capable of performing many different tasks simultaneously

- **Adaptivity**: Fine-tuning, in-context learning

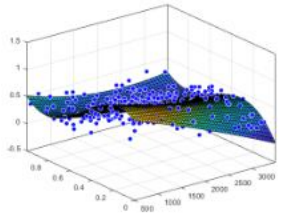
A conflux of ideas and intellectual traditions



neural AI



symbolic AI



statistical AI

- This concludes our tour of the three stories that make up what AI is today
  - Symbolic AI took a top-down approach and failed to fulfill its original promise. But it offered a vision and did built impressive artifacts for ambitious problems like QA and dialogue systems
  - Neural AI is bottom-up, starting with simple tasks. It offered a class of models, deep networks, with today's data and computing resources, capable of conquering ambitious problems
  - Statistical AI foremost offers mathematical rigor and clarity. For example, we define an objective function separate from the optimization algorithm, or have a language to talk about model complexity in learning

Expected outcomes

- You will learn the coolest AI technology, tools during the semester, and grasp the key concept for understanding their working

- During the lectures, my goal is to help you understand the principles of how these methods work, when and how we should use which methods

- **Feel free to ask questions**!

Expected outcomes

- The homework is designed to help you refresh course materials and better understand them
  - Both mathematical and coding questions are involved
  - We encourage you to form groups to complete the homework
  - Ample office hour sessions focused on debugging your coding skills

- Learn how to develop and apply state-of-the-art LLM techniques to downstream tasks

- Learn about prompt tuning, optimization, and AI evaluation

Reading materials

- Chapter 1, "AI: A modern Approach," Russell and Norvig
- McKinsey survey: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work
- Wikipedia article: https://en.wikipedia.org/wiki/History_of_artificial intelligence