Machine Learning for Biology and Health

CSCI 1851
Spring 2026

Ritambhara Singh

January 22, 2026

Thursday

ChatGPT prompt “minimalist landscape painting a cute penguin working on a computer with DNA on its screen”

About your instructor!

7th year CS and Data Science faculty and a Center for Computational Molecular Biology (CCMB) member!

Research Interests

Machine Learning, Data Mining, Computational Biology, Health Sciences

B.E
208-2012

Ph.D.
2012-2018

Postdoc
2018-2019

Office Location

Room 313, Data Science Institute (DSI)
3rd Floor, 164 Angell Street

Office Hours (Starting this week!)

Fridays, 3:30-5:30 PM (using appointment slot) ← May move around

Where: Room 313, 3rd Floor, 164 Angell Street ← Not CIT!

Email: ritambhara@brown.edu

Website: www.ritambharasingh.com

Introducing...

Your Awesome Course Staff!

Your HTAs!

Falak Pabari
(Senior)

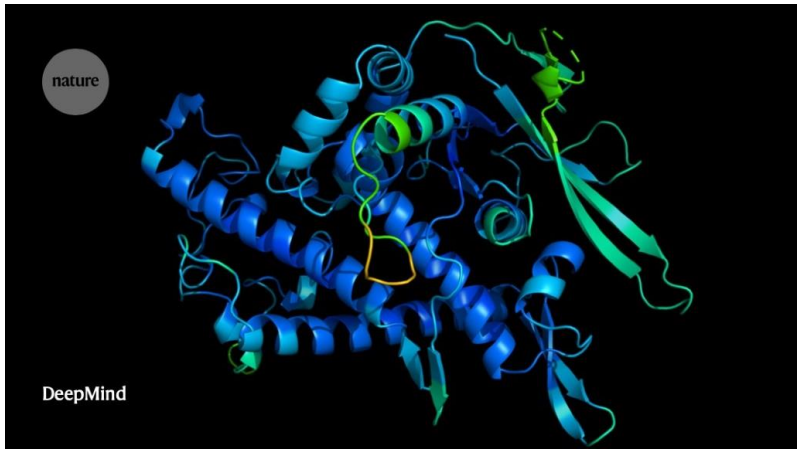
Kyle Yeh
(Senior)

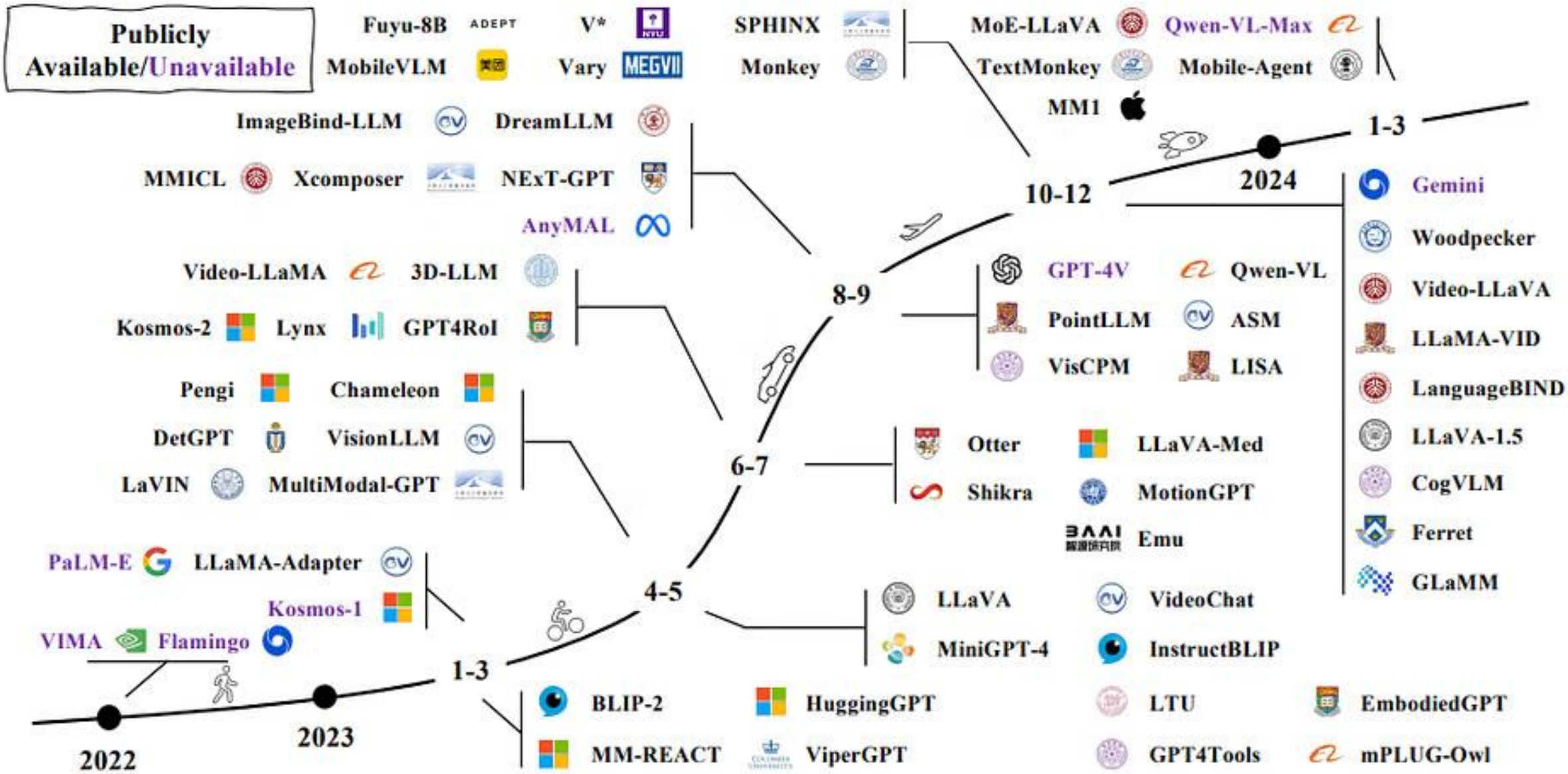
Course Advisor

Ioanna Gemou
(Ph.D. student)

Why take this course?

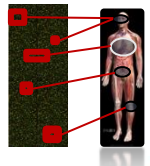
You may have heard of “Artificial Intelligence (AI)”




Publicly Available/Unavailable
2022
2023
2024
1-3
4-5
6-7
8-9
10-12
PaLM-E
LLaMA-Adapter
Kosmos-1
VIMA
Flamingo
BLIP-2
MM-REACT
HuggingGPT
ViperGPT
LLaVA
MiniGPT-4
VideoChat
InstructBLIP
LTU
GPT4Tools
EmbodiedGPT
mPLUG-Owl
Otter
Shikra
LLaVA-Med
MotionGPT
Emu
Pengi
Chameleon
DetGPT
VisionLLM
LaVIN
MultiModal-GPT
Video-LLaMA
3D-LLM
Kosmos-2
Lynx
GPT4RoI
ImageBind-LLM
DreamLLM
MMICL
Xcomposer
NExT-GPT
AnyMAL
GPT-4V
Qwen-VL
PointLLM
ASM
VisCPM
LISA
Fuyu-8B
ADEPT
V*
MobileVLM
Vary
SPHINX
Monkey
MoE-LLaVA
Qwen-VL-Max
TextMonkey
Mobile-Agent
MM1
Gemini
Woodpecker
Video-LLaVA
LLaMA-VID
LanguageBIND
LLaVA-1.5
CogVLM
Ferret
GLaMM

Health is the next frontier!!!

Molecular signatures of tumor / blood sample



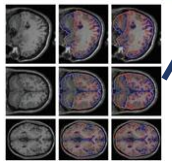
Signs & Symptoms

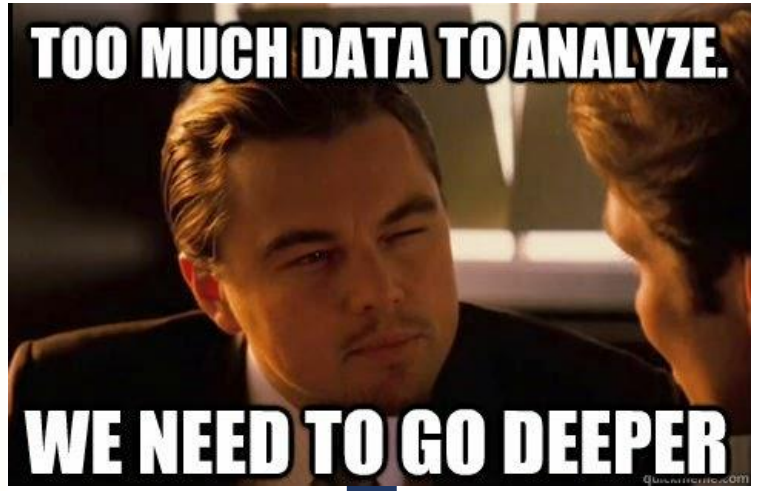
Patient Medical History & Demographics

Public Health Data

Genetic Data

Mobile medical sensor data





Artificial Intelligence

ChatGPT has some thoughts…

3 reasons to take the course "machine learning for biology and health"

Here are **3 strong reasons** to take the course **"Machine Learning for Biology and Health"**:

1. **Tackle real-world biological and medical problems**
   You'll learn how machine learning is applied to genomics, medical imaging, drug discovery, and disease prediction—skills that directly impact healthcare and life sciences rather than abstract toy problems.
2. **Gain in-demand interdisciplinary skills**
   Combining ML with biology and health makes you highly competitive. Employers and research labs increasingly seek people who can bridge data science with biological insight, a skill set that's still relatively rare.
3. **Prepare for research, industry, and innovation**
   The course builds a foundation for cutting-edge research and practical applications, helping you read scientific papers, design experiments, and develop ML-driven solutions in biotech, pharma, and digital health.

https://openai.com/blog/chatgpt/

Our goal is to answers some important questions

- What is are the different machine learning and deep learning models?
- How to implement these models for biology and health applications?
- What models are appropriate for different applications?
- Will our approach improve our understanding of the data or the problem?

DALL-E 2 prompt “a cartoon of a penguin with a question mark”

AI or Machine Learning or Deep Learning?

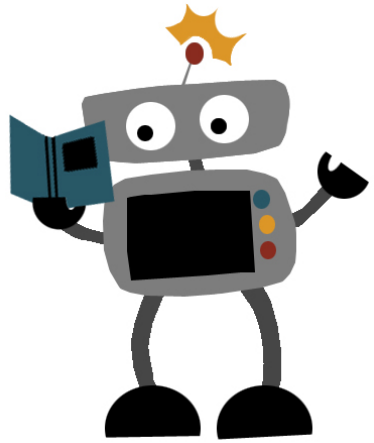
What is the distinction?

(1) What is Machine Learning?

(2) How does it connect to Deep Learning?

(3) What is AI?

What is Machine Learning?



Input: X

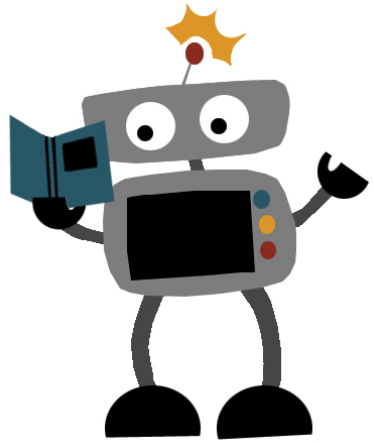
Function: f

f(X)→Y

Output: Y

”Cooking?”

What is Machine Learning?



Supervised Learning

Input: X

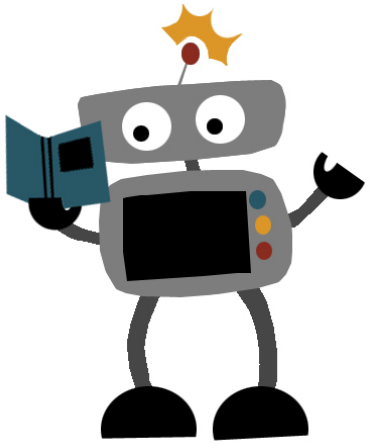
Learned function: f

Output: Y

"Cooking?"

$f(X) \rightarrow Y$

What is Machine Learning?



Input: X

I do not want sour cream in my burrito

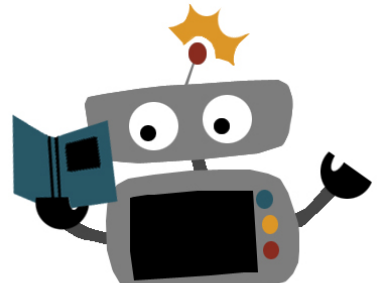
Learned function: f

Output: Y

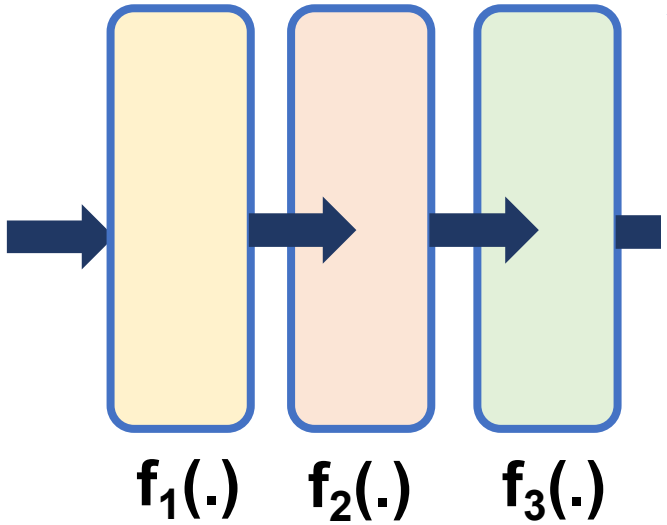
No quiero crema agrea en mi burrito

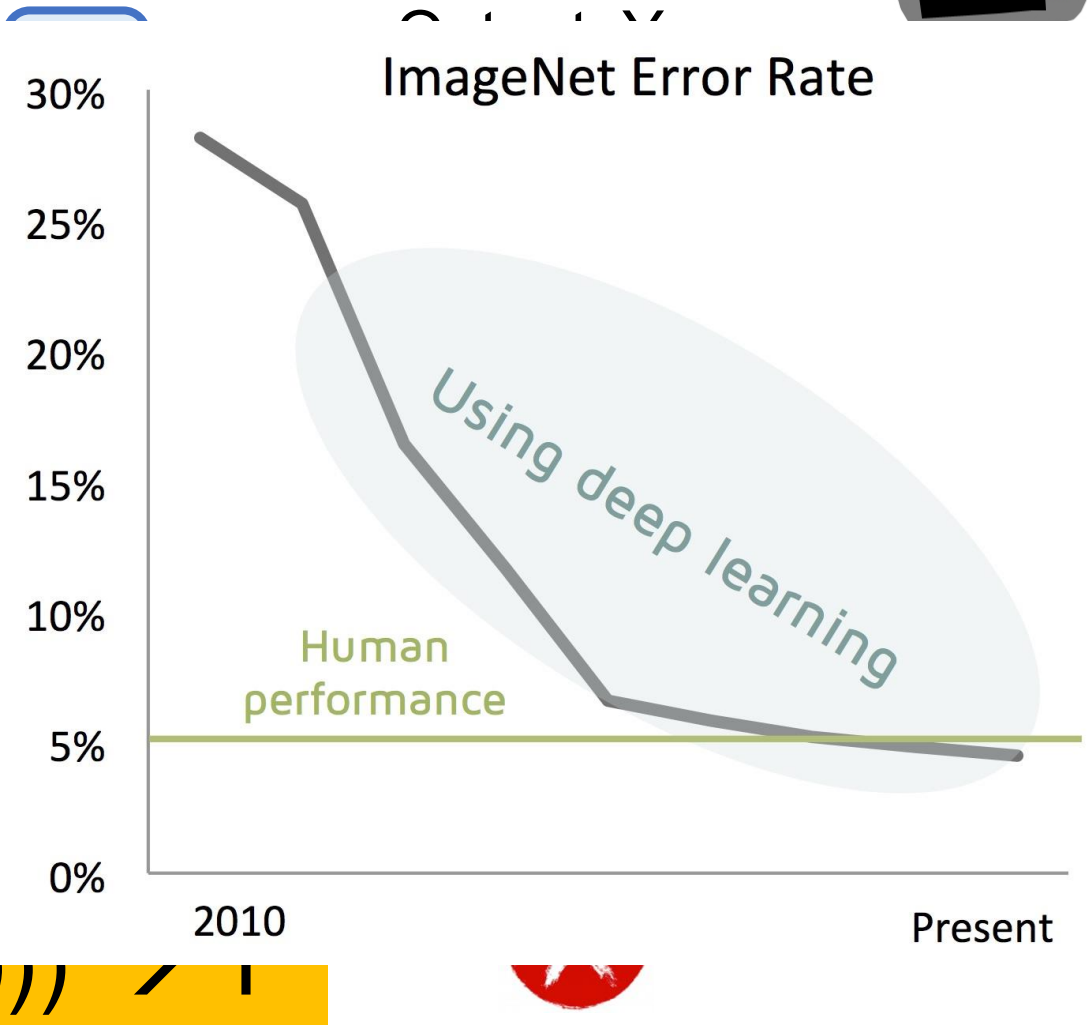
$f(X) \rightarrow Y$

What is Deep Learning?



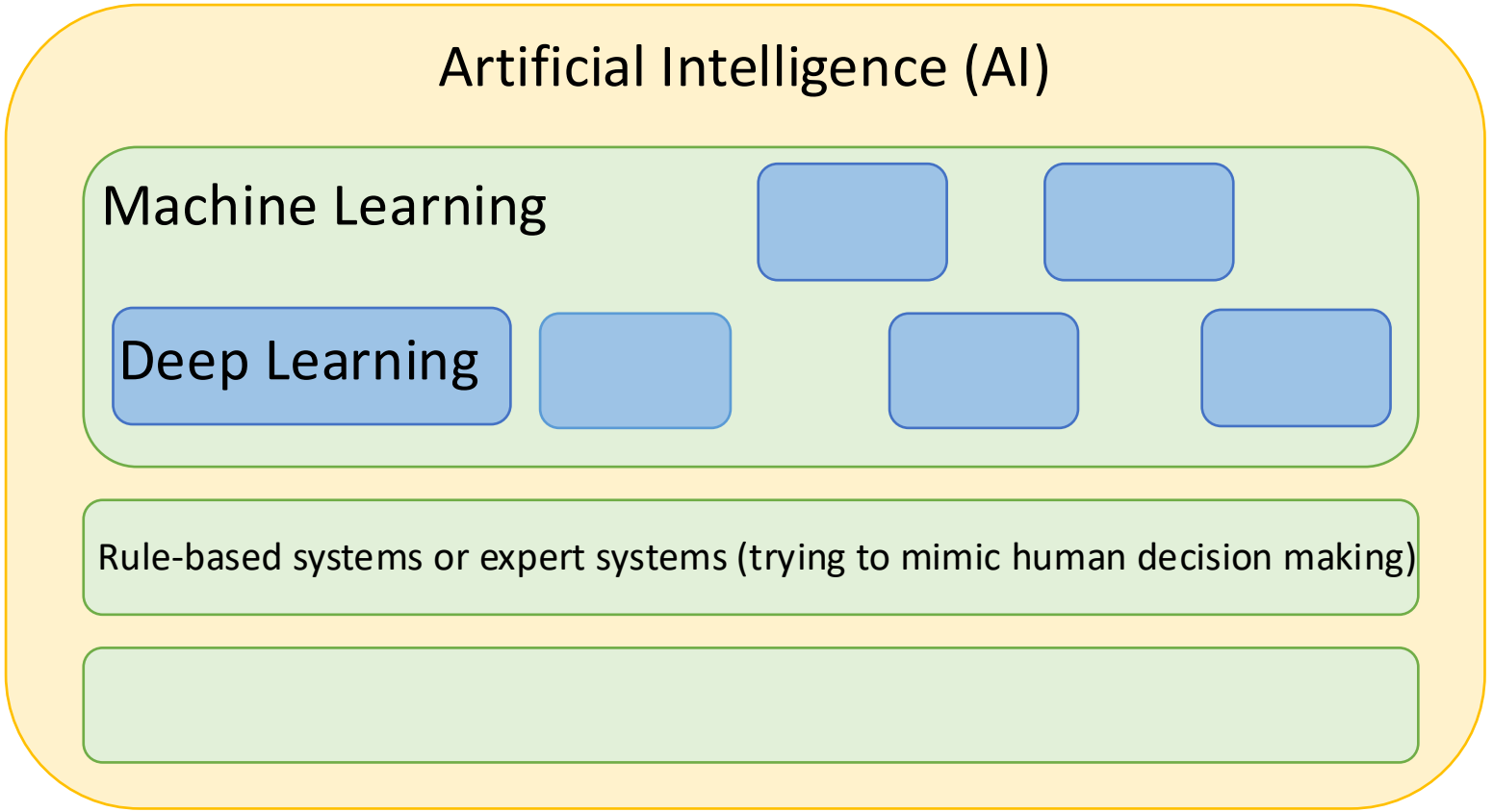
Input: X



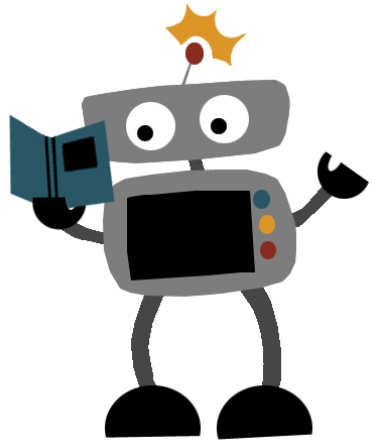


$f_4\,(f_3\,(f_2\,(f_1\,(X))))$

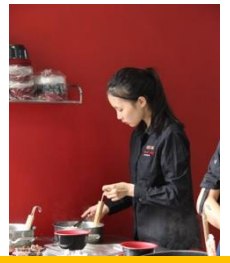
What is AI?


Artificial Intelligence (AI)
Machine Learning
Deep Learning
Rule-based systems or expert systems (trying to mimic human decision making)

Recap



Input: X



Machine Learning

$f(X) \rightarrow Y$

Output: Y

"Cooking?"

AI, machine learning, deep learning are connected but not the same

Deep Learning

$f_4\,(f_3\,(f_2\,(f_1\,(X)))) \rightarrow Y$

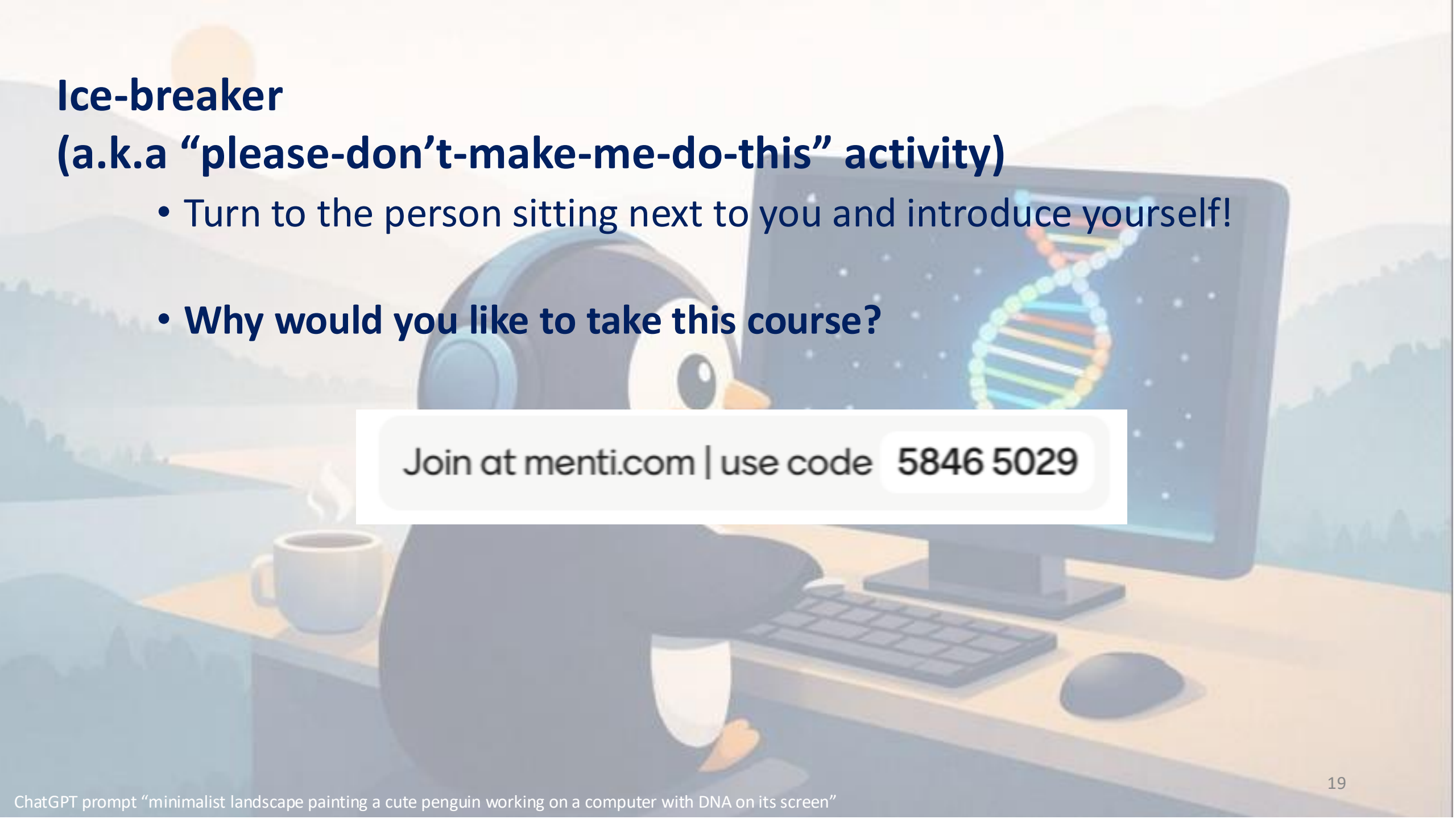
Questions?

Ice-breaker
(a.k.a “please-don’t-make-me-do-this” activity)

- Turn to the person sitting next to you and introduce yourself!

- **Why would you like to take this course?**

Join at menti.com | use code 5846 5029



ChatGPT prompt “minimalist landscape painting a cute penguin working on a computer with DNA on its screen”

Course Logistics

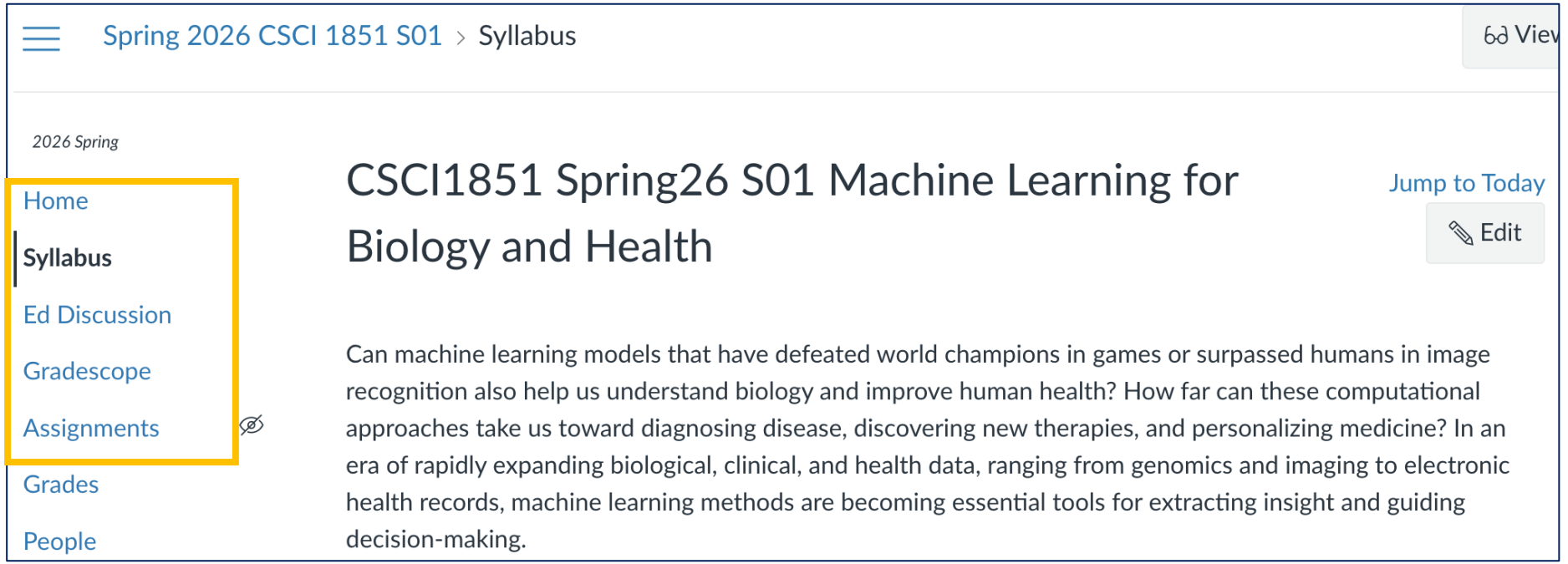
The Course Website

https://csci1851.github.io/

- Your one-stop-shop for:
  - Syllabus
  - Lecture & assignment schedules
  - Links to important forms, etc.
  - …

The Canvas Website

https://canvas.brown.edu/courses/1102089



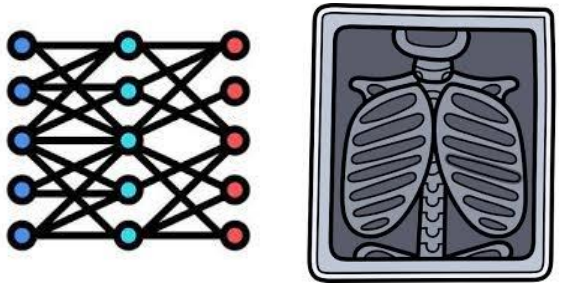
- Your access to:
  - Ed Discussion
  - GradeScope
  - Weekly course announcements from instructor

Six Awesome Assignments

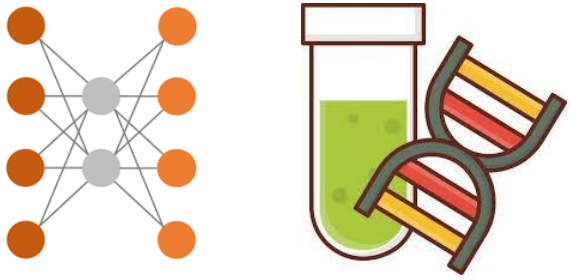
Linear models

Decision trees

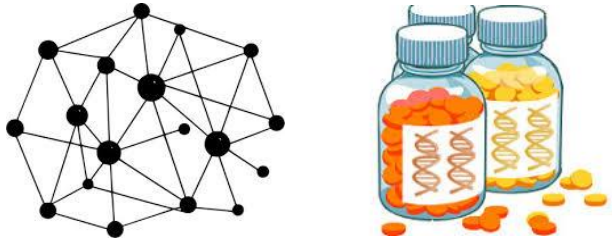
Support Vector Machines



Neural Networks and Convolutional Neural Networks



Autoencoders and Variational Autoencoders



Graph Neural Networks

Assignment logistics

- Homework Assignments
  - Get stencils via Github
  - Submission via Gradescope

NOTE:
Minimal stencil code, but with clear instructions (think working in a research lab!)

Can look up functions, information – PLEASE do not ask AI to code for you

Expect coding related questions in the exam

Two exams (in person)

Mid-term Exam (75mins)

- Scheduled for March 19, 2016
- Will focus on Sections 1-3

Final Exam (Duration TBD)

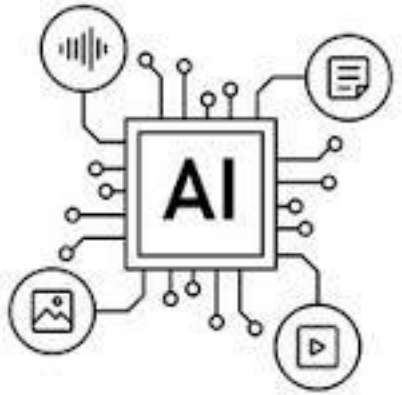
- Scheduled for May 08, 2026
- Will focus on Section 3-6, with 1-2 questions from Sections 1-3

Final course project!

- Form a team

Round 1 (Mid-term)

- We will release the details of the prediction task and datasets
- Apply machine learning model to solve the task
- **Mid-term**: Submit trained model
- We will release score board on held out dataset



Multimodal AI

kaggle

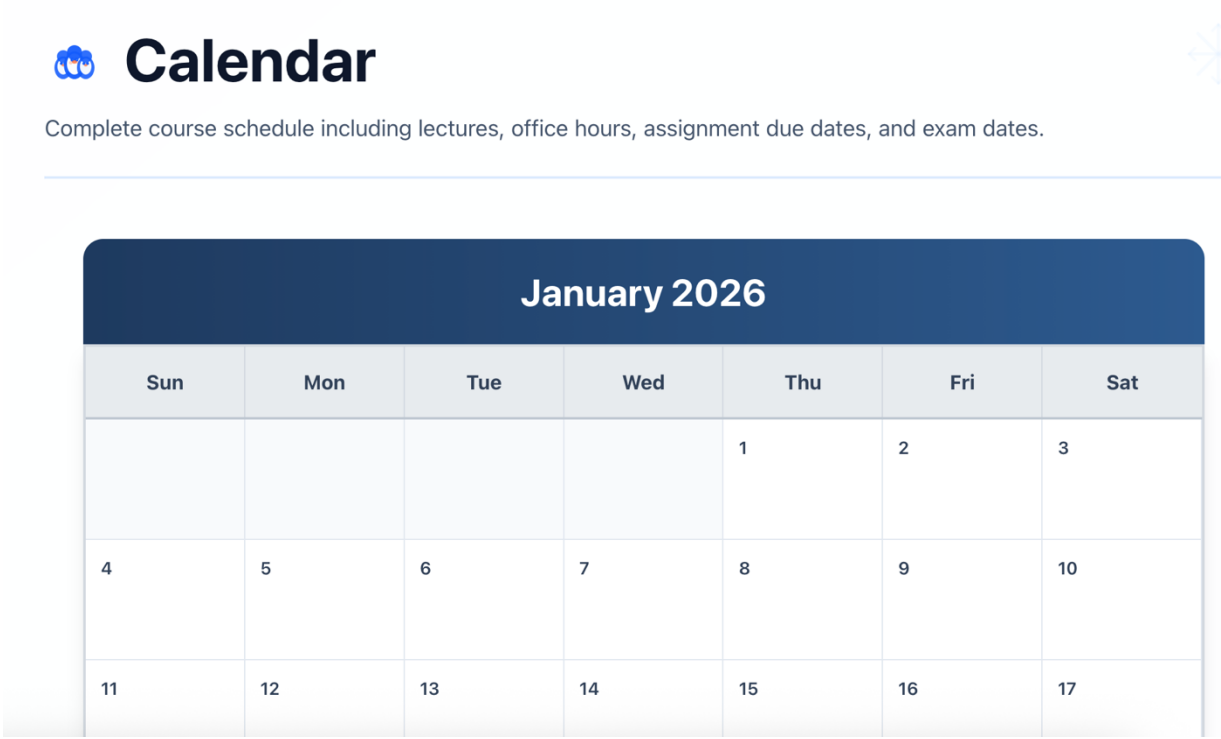
Round 2 (Final)

- We will update the task and introduce new datasets
- Utilizing previous results + new data, update the model for the task
- **Final:** Submit trained model, present the work (May 05), and write a report (May 12)
- We will release score board on held-out dataset

HTA office hours

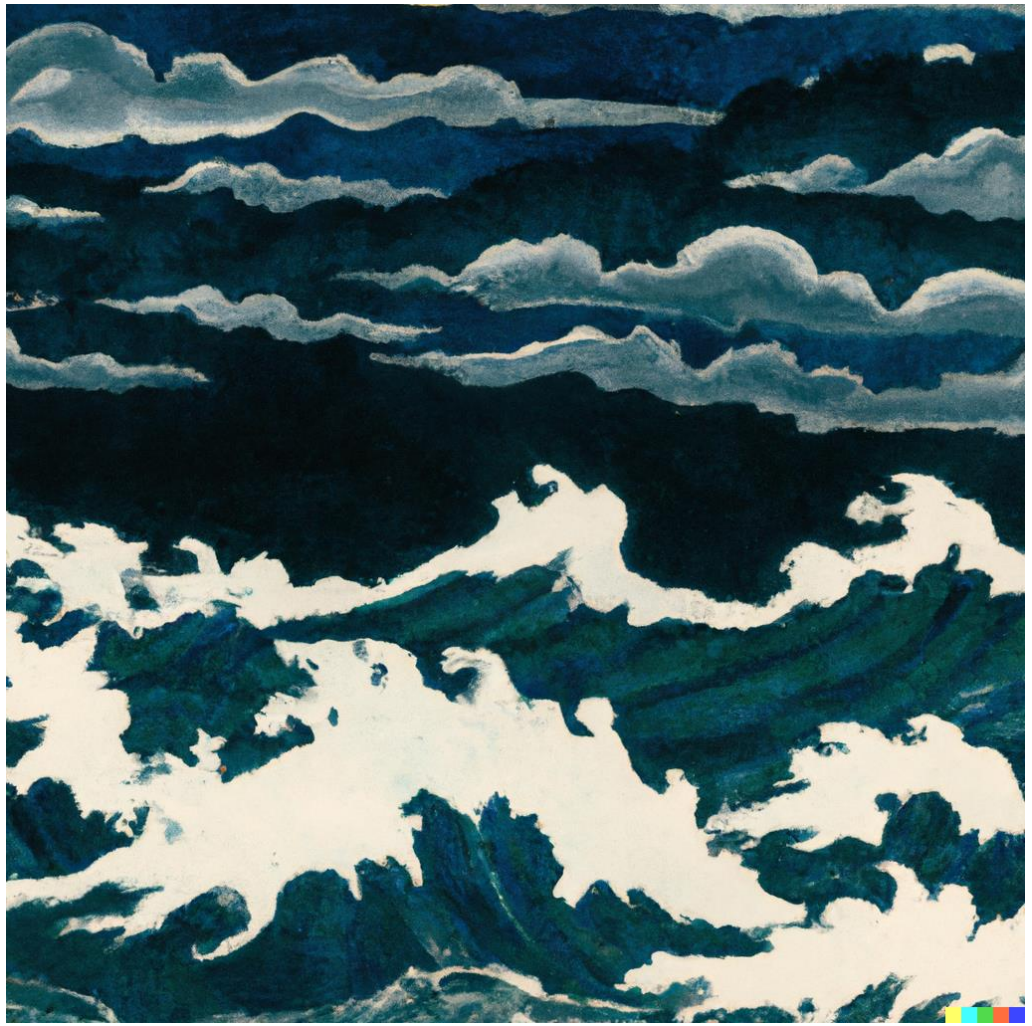
- HTA office hours (in-person)
  - Will take a week or so to finalize
  - Might have remote options (in the works!)

- **FORMAT** for HTA office hours
  - Conceptual/Collab hours
    - group-based help/discussion on concepts)
    - group-based help/discussion on code



The only thing set in stone is our mission to learn!

- ❑ First offering of this course!
- ❑ Will try things for class engagement
- ❑ Due dates might move around
- ❑ We might have make-up classes
- ❑ The schedule will remain flexible till the end
- ❑ Suggestions are welcome!
  Patience is appreciated!



DALL-E 2 prompt “japanese painting of a stormy ocean”

Acknowledgements
(Course development and testing)

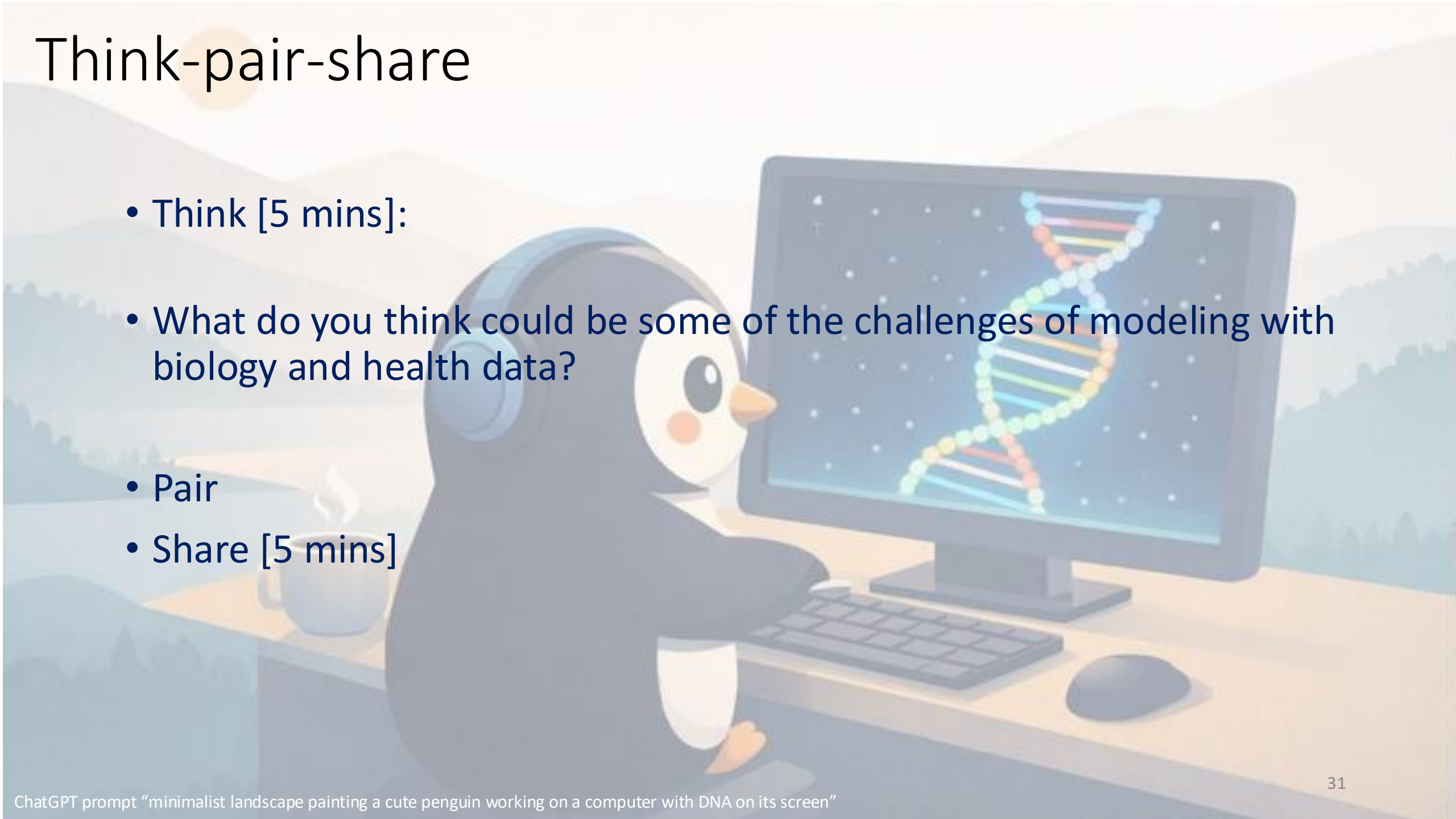
Pranav Mahableshwarkar

Jason Lin

Questions?

Think-pair-share

- Think [5 mins]:
- What do you think could be some of the challenges of modeling with biology and health data?
- Pair
- Share [5 mins]



ChatGPT prompt “minimalist landscape painting a cute penguin working on a computer with DNA on its screen”

Immediate Action Items

- **Read the course syllabus**
- Make sure you can access the course Ed Discussion page via Canvas (all announcements will happen there)
- Create a GitHub account (if you don't have one already)
- Make sure you can access the course GradeScope via Canvas (all assignments will be submitted there)
- If you want to do work on CS department machines over SSH, you'll need a CS login
  - All enrolled students automatically have one (it's the same as your Brown ID)
  - If you are not yet enrolled, you'll need to email problem@cs.brown.edu
  - More information about CS accounts can be found here.
  - *Note that a CS login is **not required** for this course, as we are not using handin or any other department infrastructure to manage files or submissions*

Explore new and existing community resources and mental health support:

BROWN

The Brown community is resilient, caring and strong.

We are ever true.