DDA4220 · Fall 2026

Lecture 1

Deep Learning and Applications

Course Introduction

Zhen Liu

CUHK-Shenzhen

The Turing test

Alan Turing, 1950

- A judge exchanges written messages
- One participant is human; one is a machine
- The judge tries to distinguish them

Symbolic AI

Facts

Alice is a parent of Ben.

Ben is a parent of Chen.

Rule

If X is a parent of Y,
and Y is a parent of Z,
then X is a grandparent of Z.

Conclusion

Alice is a grandparent of Chen.

Writing a spam filter

If the subject contains “free”, mark it as spam.

FREE prize — claim now!	Unwanted promotion
Are you free after class?	Ordinary conversation
Your exclusive reward is waiting	Unwanted promotion

The same keywords in different messages

Scam message

"Claim your free prize now."

Campus announcement

"Claim your free student ticket."

Ambiguity in machine translation

“I saw her duck.”

我看见了她的鸭子。

我看见她低头躲开。

Early machine translation: unmet expectations

LANGUAGE AND MACHINES

COMPUTERS IN TRANSLATION AND LINGUISTICS

A Report by the
Automatic Language Processing Advisory Committee
Division of Behavioral Sciences
National Academy of Sciences
National Research Council

Publication 1416

National Academy of Sciences — National Research Council

Washington, D. C. 1966

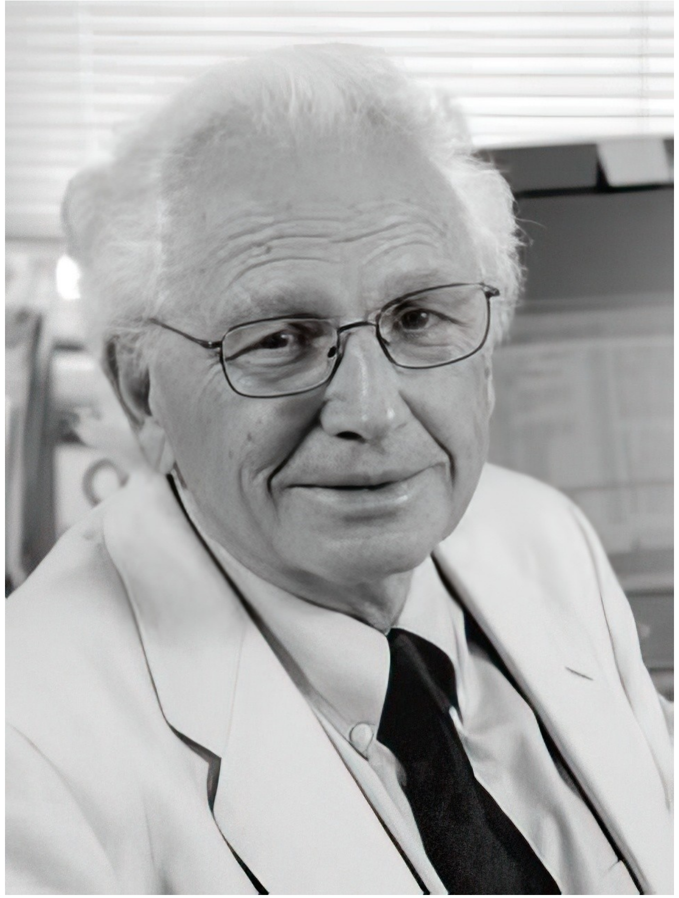
1954 · Georgetown–IBM

A controlled demonstration
250 words · 6 grammar rules

1966 · ALPAC report

- Poor translations and costly human revision
- US research funding fell sharply

Statistical language processing



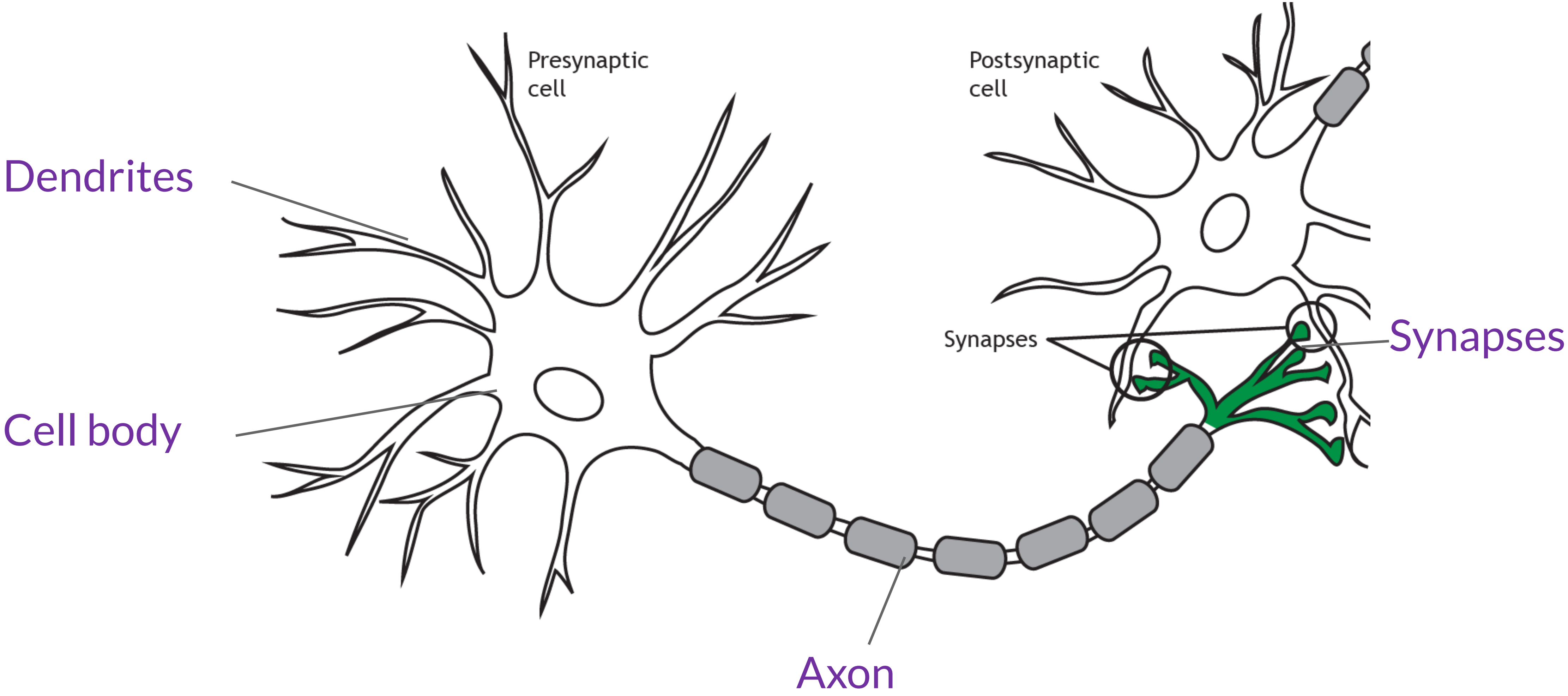
Frederick Jelinek

IBM · Speech recognition

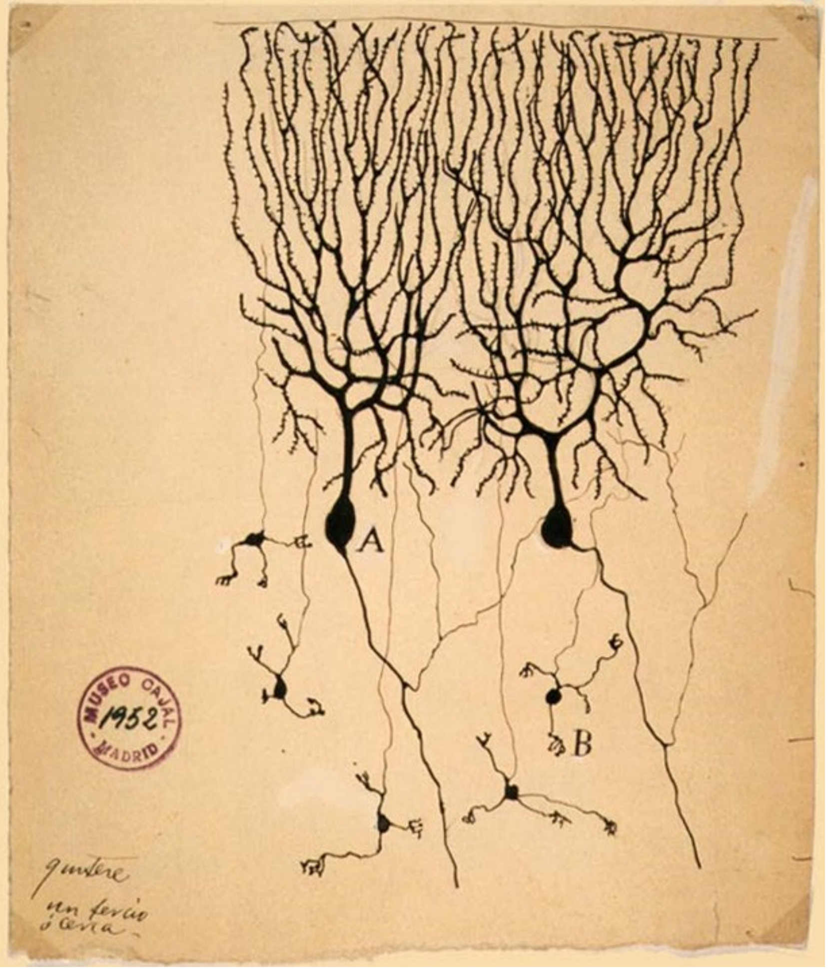
“Whenever I fire a linguist
our system performance improves”

Models learned language patterns
from text.

Biological neurons



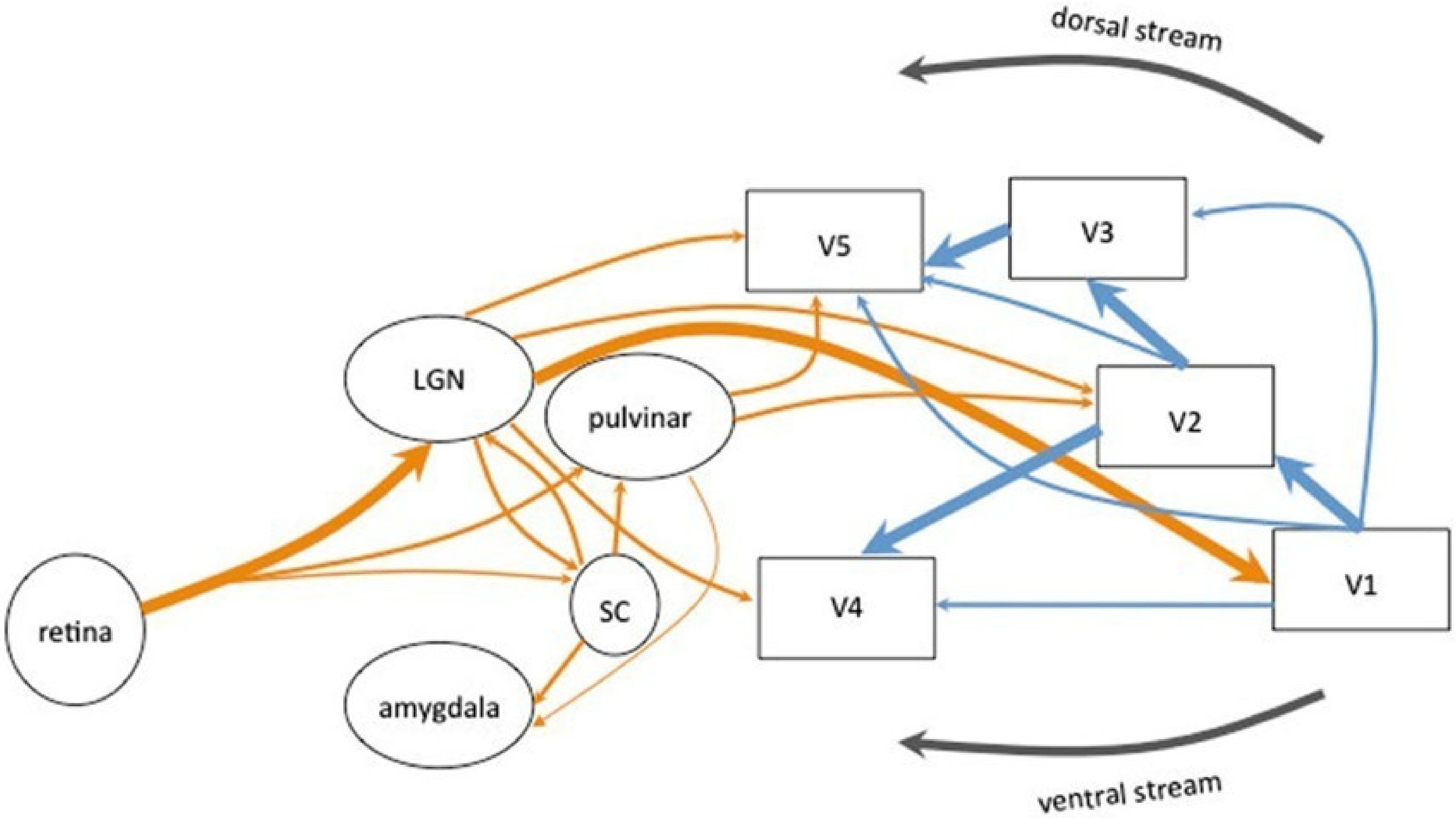
Connectionism



A network of interacting cells

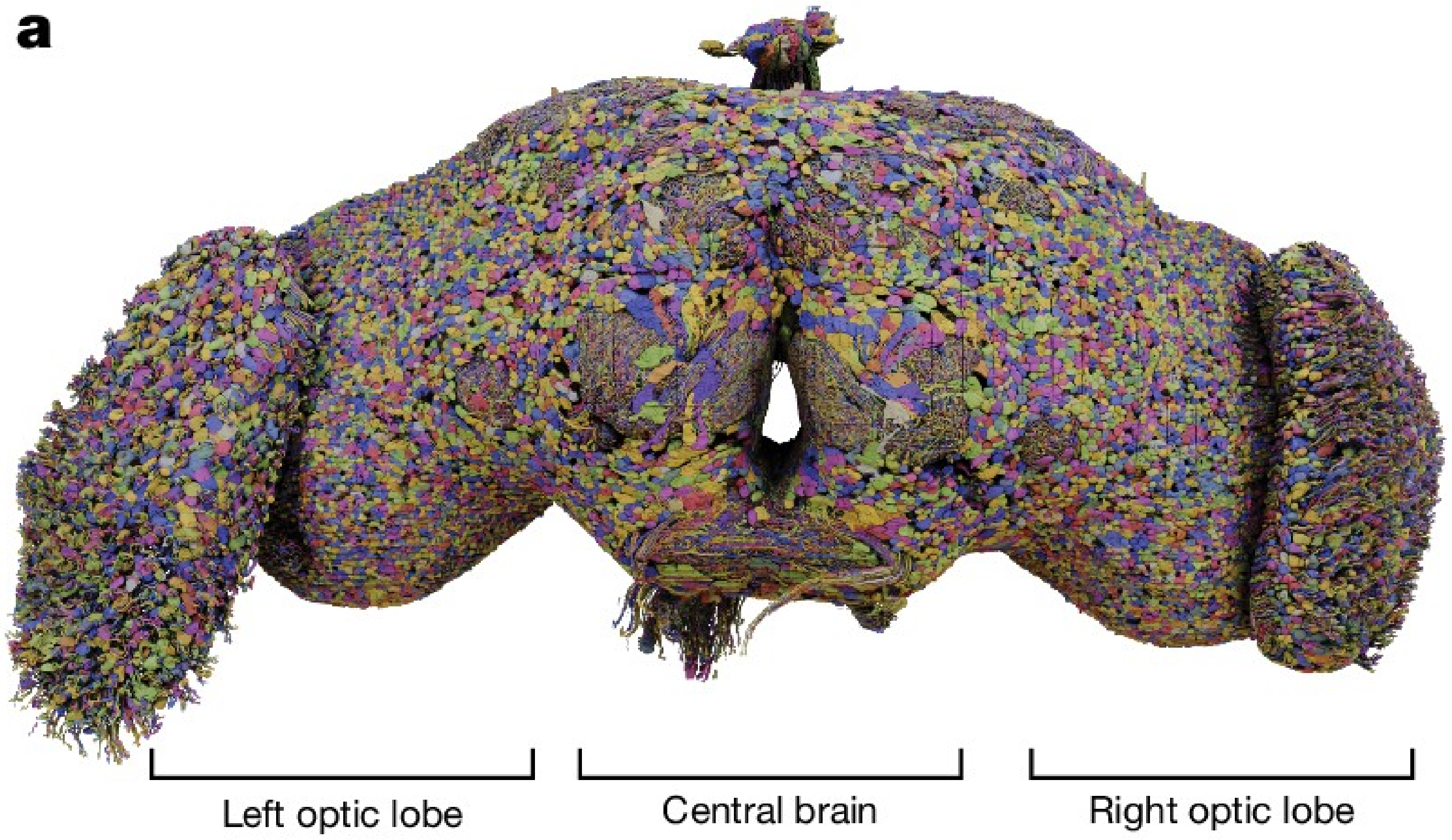
- Many units contribute to a response
- Connections influence the computation
- Learning changes the connections

The human visual cortex



A fruit-fly brain

Adult female fruit fly · 2024 reconstruction



139,255

neurons

≈50 million

synapses

Human brain

≈86 billion neurons

Early connectionist ideas

1943 **McCulloch and Pitts**

A simple computational neuron

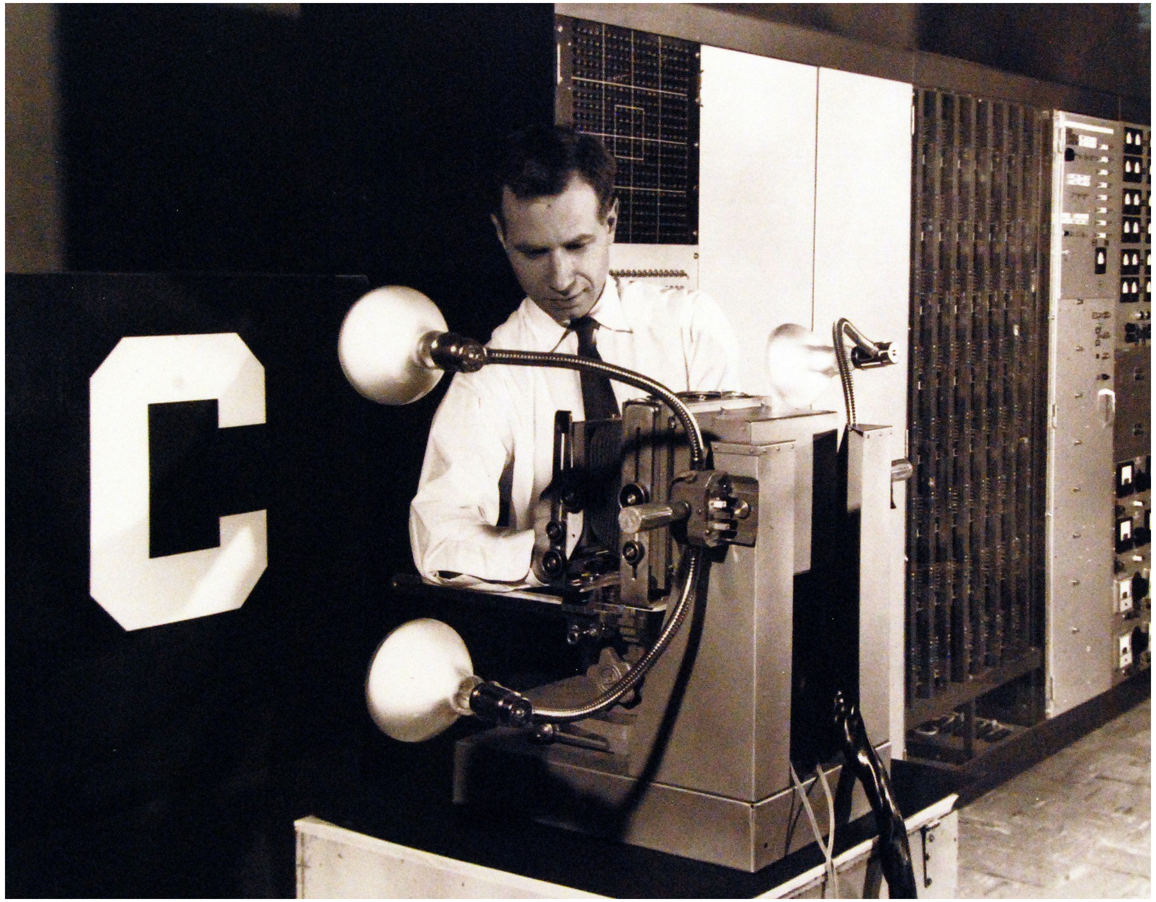
1949 **Donald Hebb**

Learning through changing connections

1958 **Frank Rosenblatt**

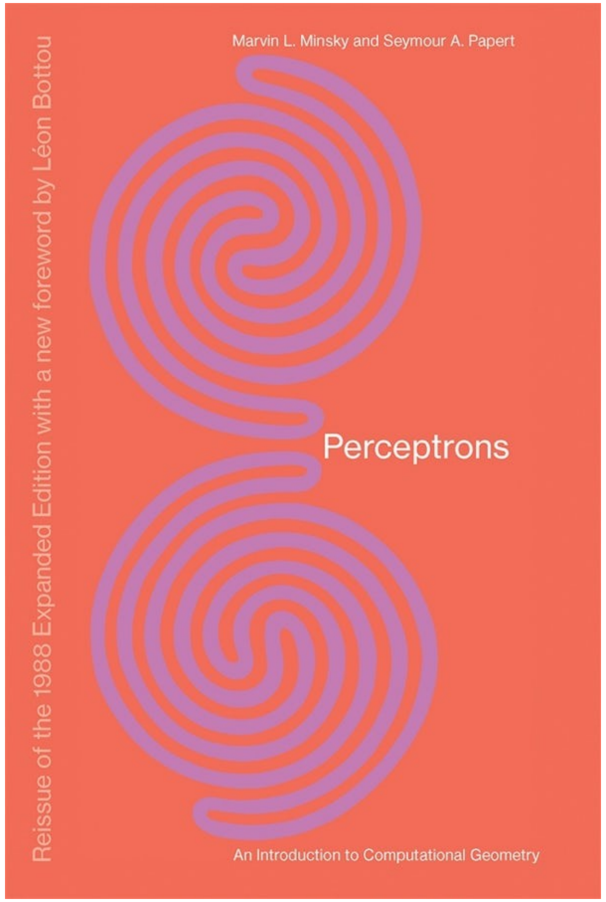
The perceptron learns from examples

The perceptron



Mark I perceptron

1969–1970s: early setbacks

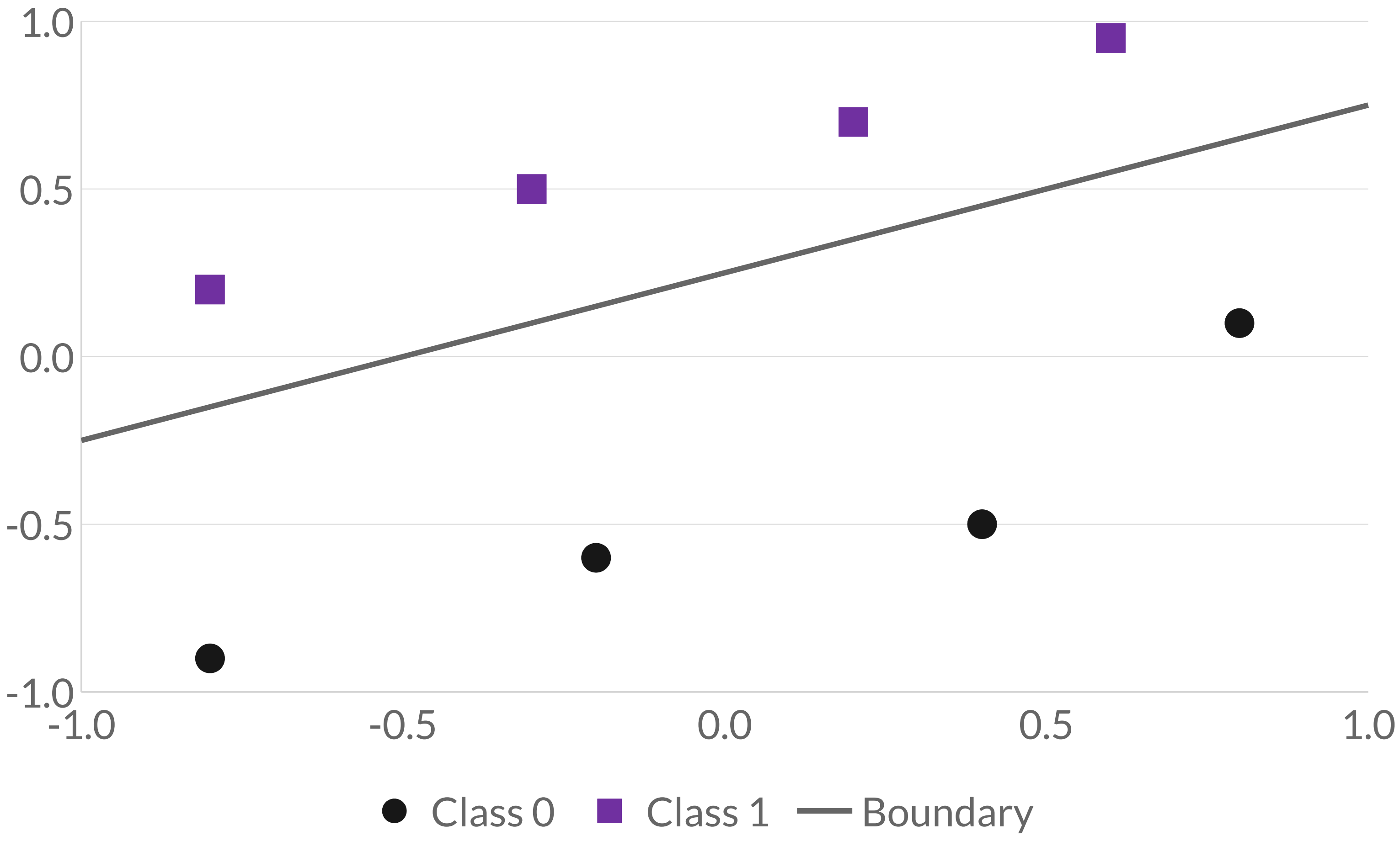


Minsky and Papert

Perceptrons, first published in 1969

- Limits of early perceptrons
- Learning larger networks remained difficult
- Expectations exceeded practical results

A linear classifier



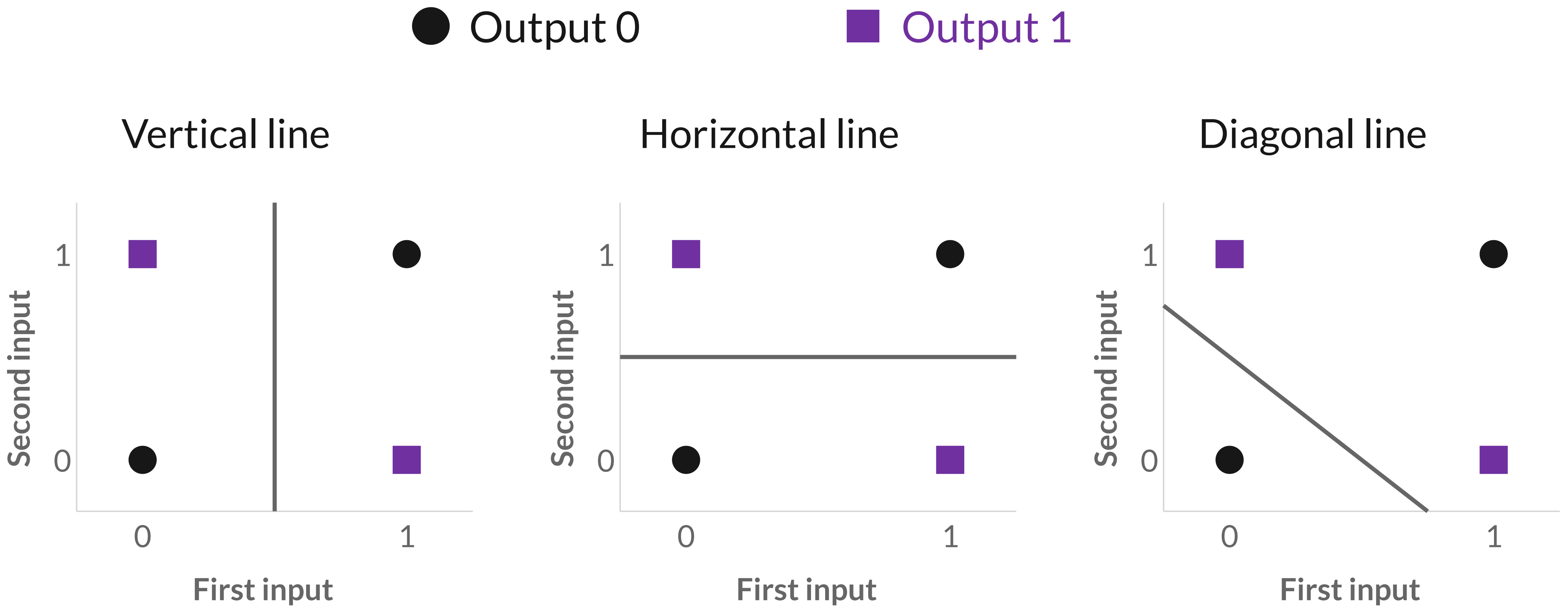
A straight line separates these classes.

XOR: exclusive OR

Output 1 when exactly one input is 1

0, 0	Output 0	0, 1	Output 1
Neither input is 1		Only the second is 1	
1, 0	Output 1	1, 1	Output 0
Only the first is 1		Both inputs are 1	

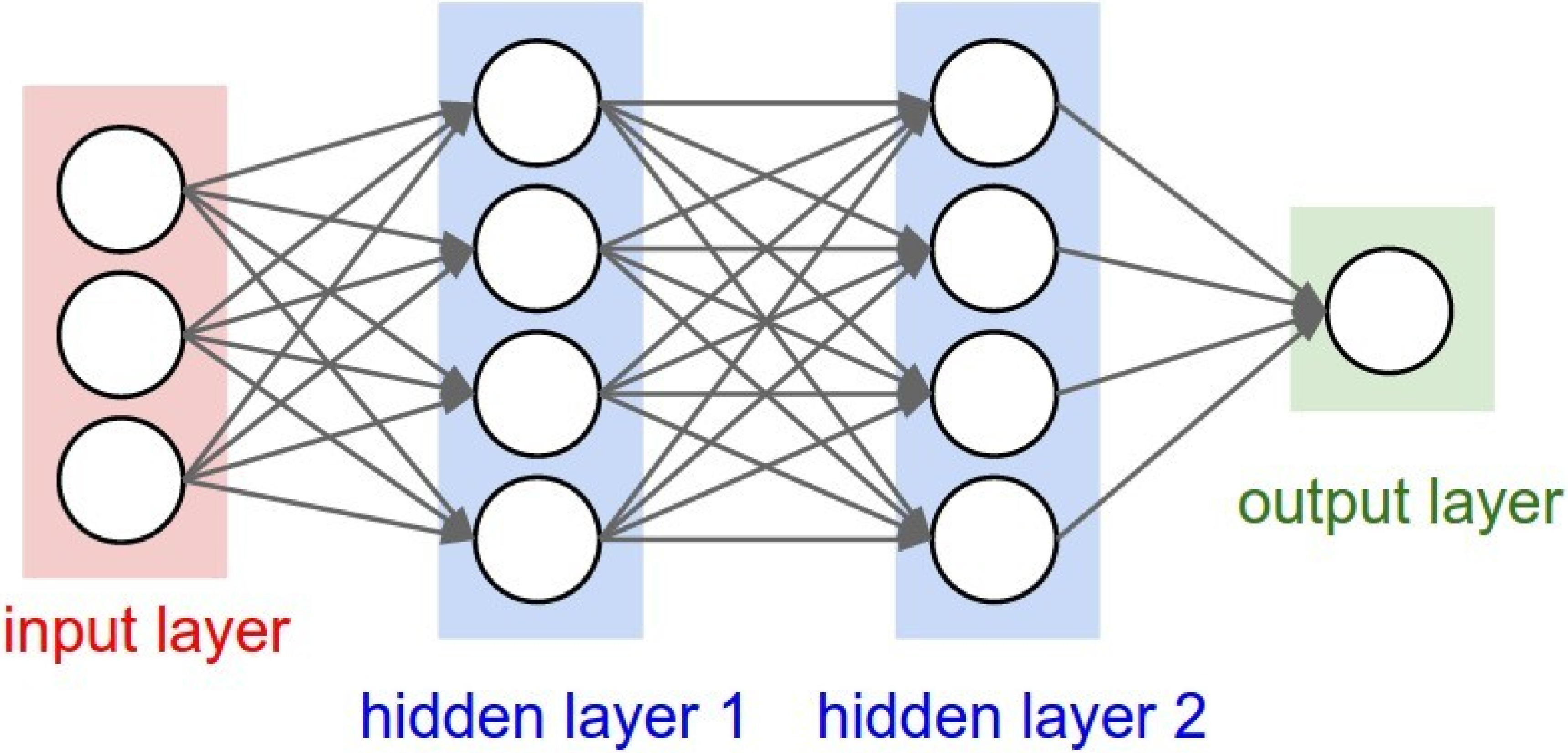
A perceptron cannot solve XOR



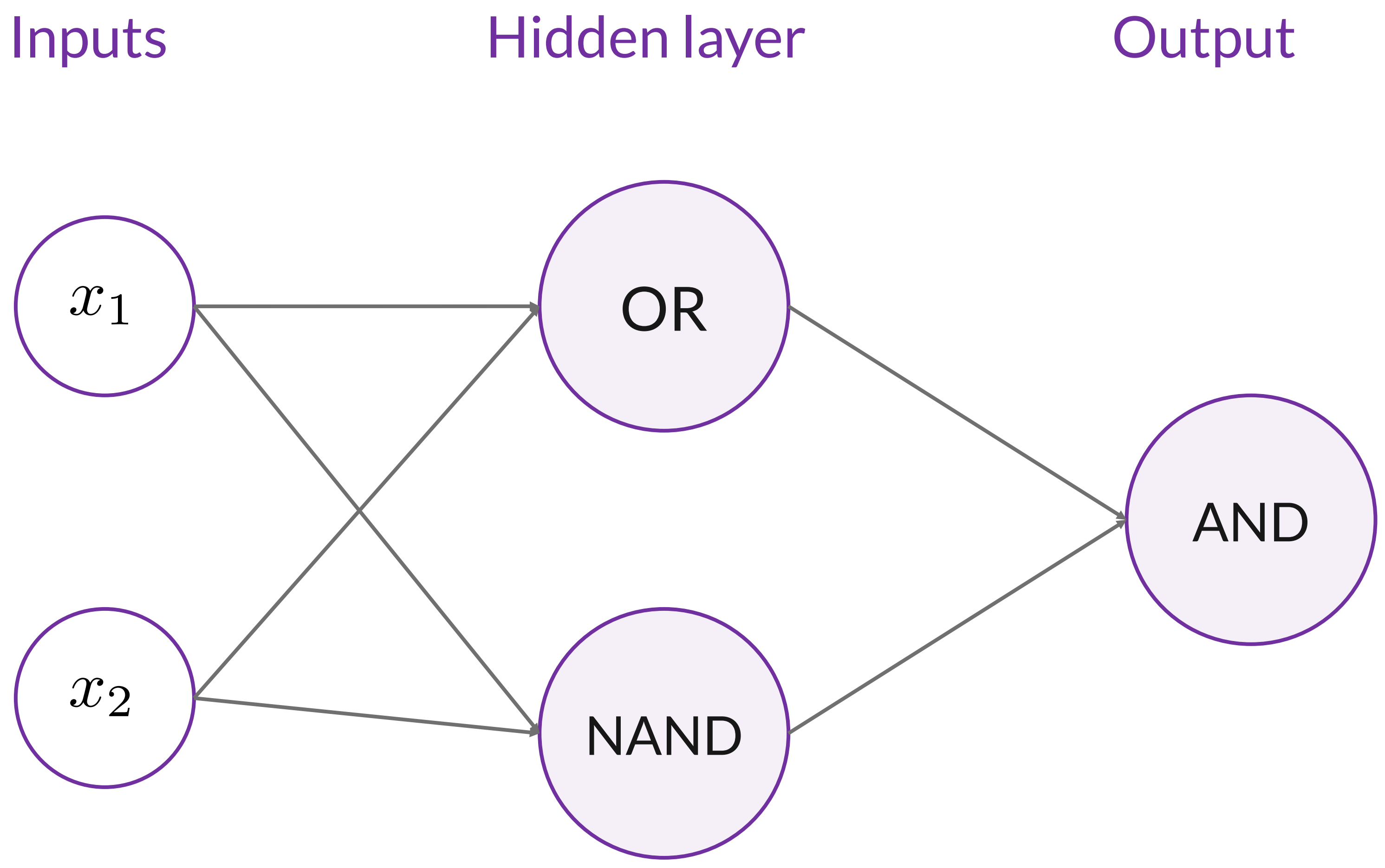
Each attempt leaves different outputs on the same side.

Networks with several layers

Multilayer perceptron (MLP)



An MLP can solve XOR



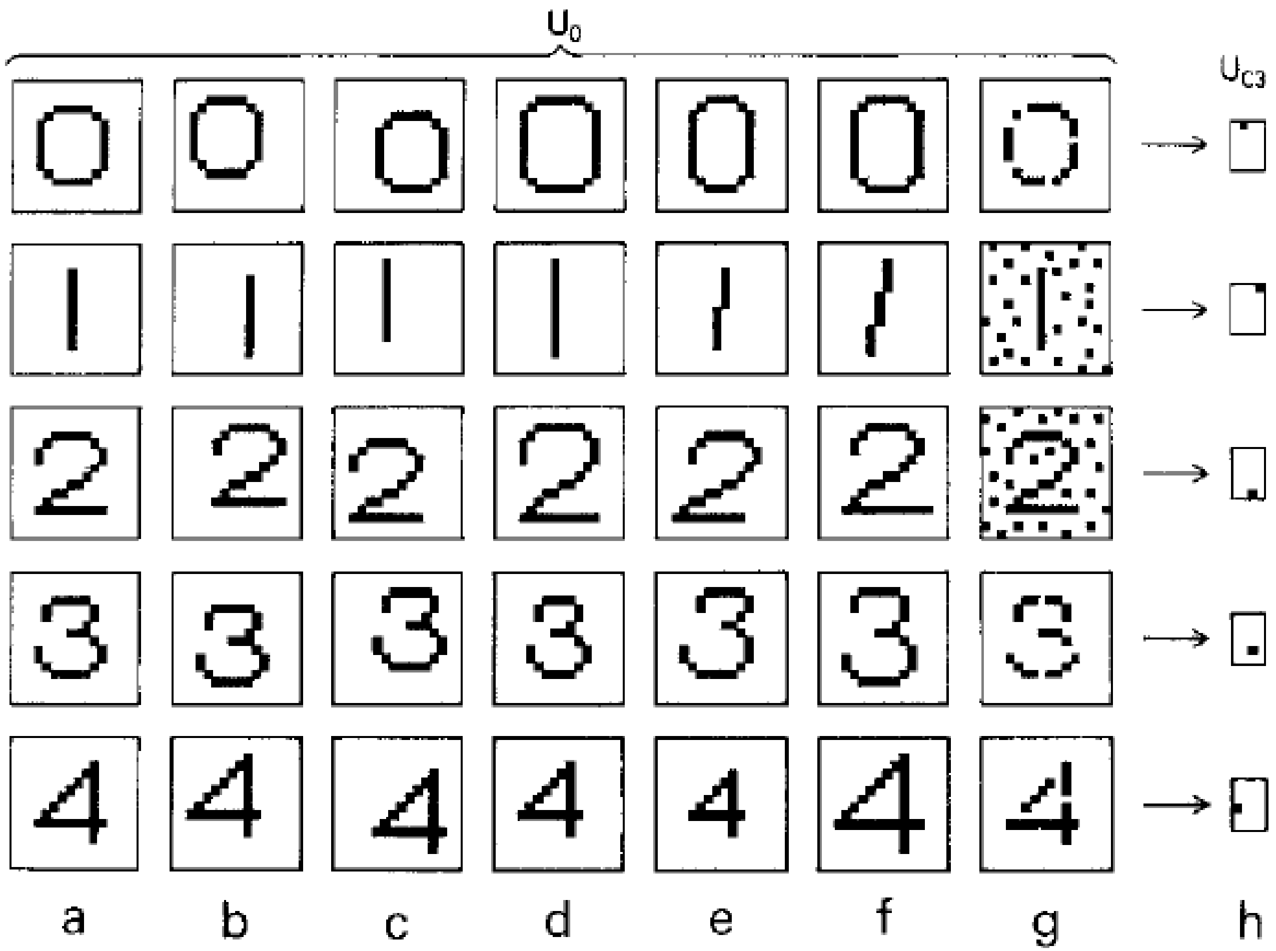
OR

At least one input is 1

NAND

Not both inputs are 1

Early CNN ideas: the neocognitron

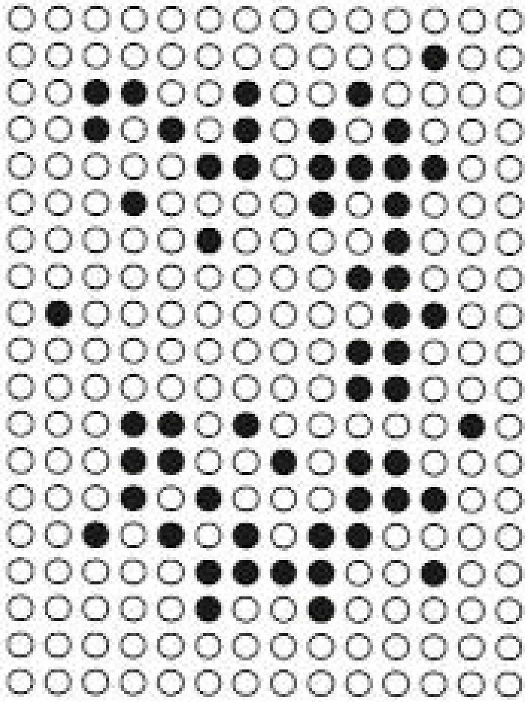


1980

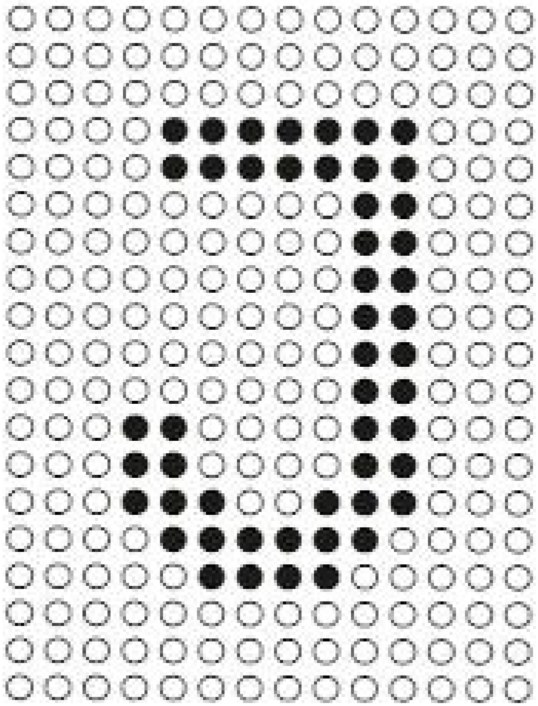
Recognizing patterns when position or shape changes

1982: Hopfield networks and associative memory

Incomplete cue

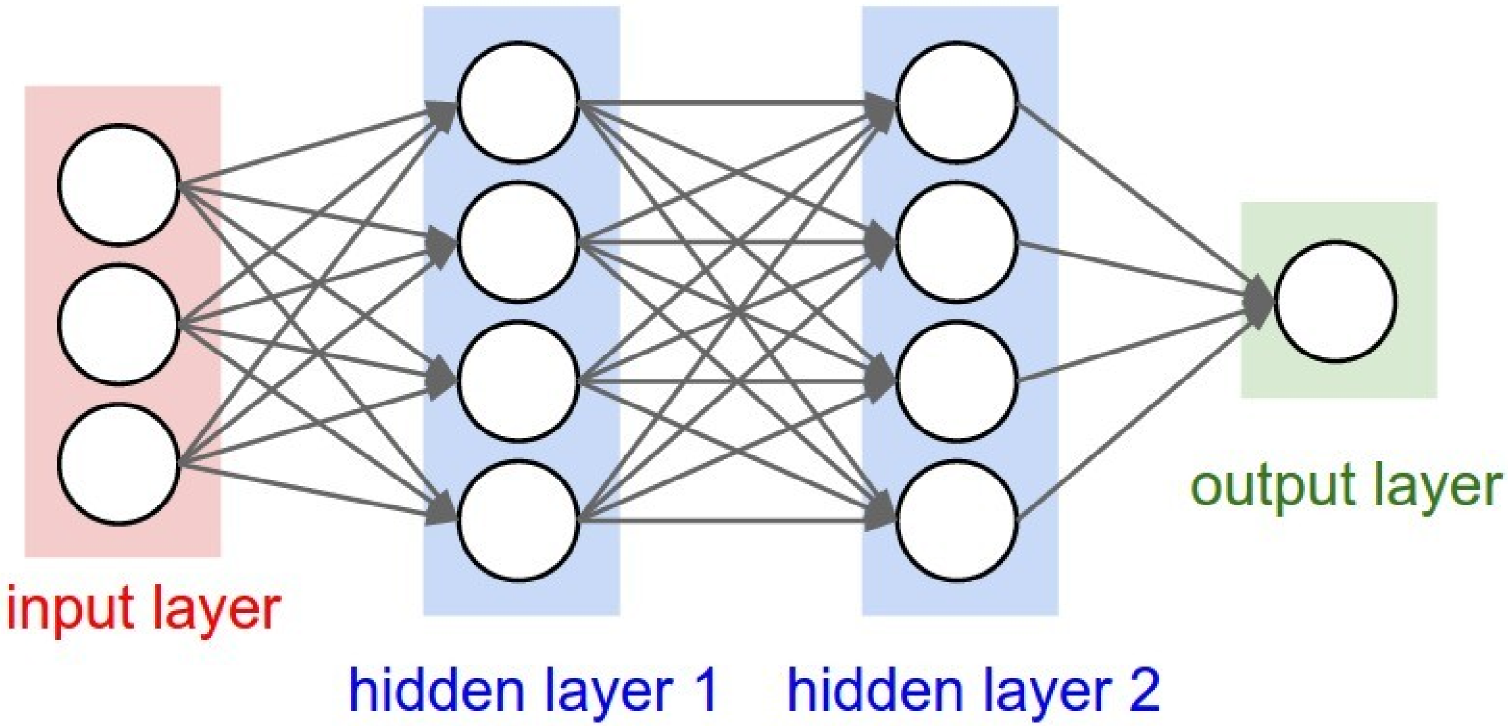


Stored pattern



Backpropagation

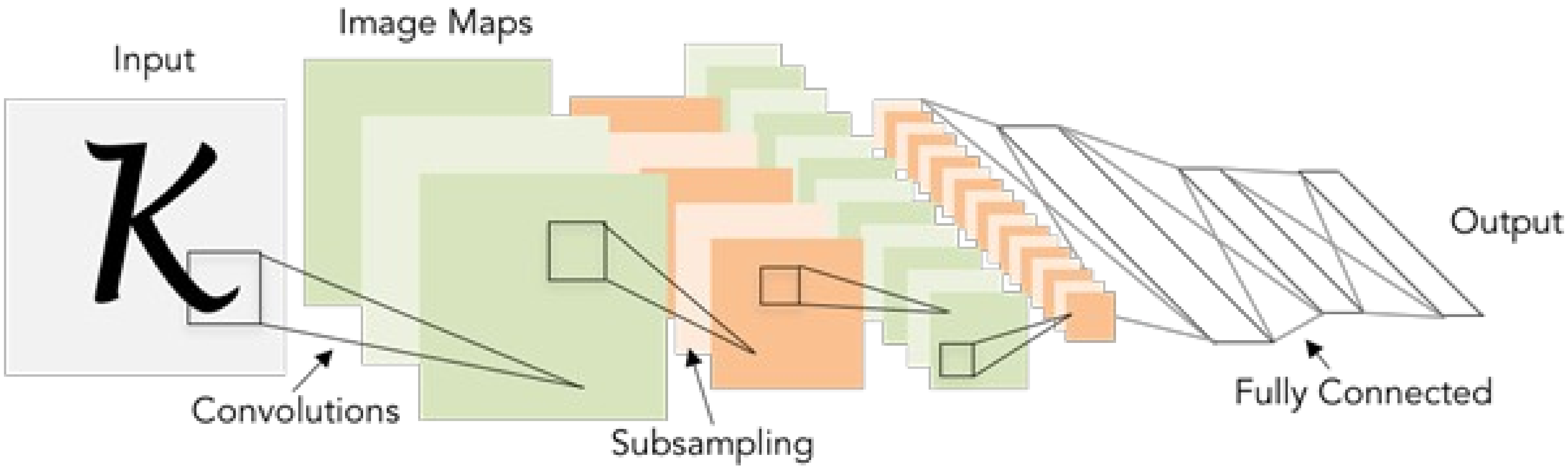
1986 · Rumelhart, Hinton and Williams



A target guides the output.

Hidden units learn without individual target labels.

CNNs become practical: LeNet

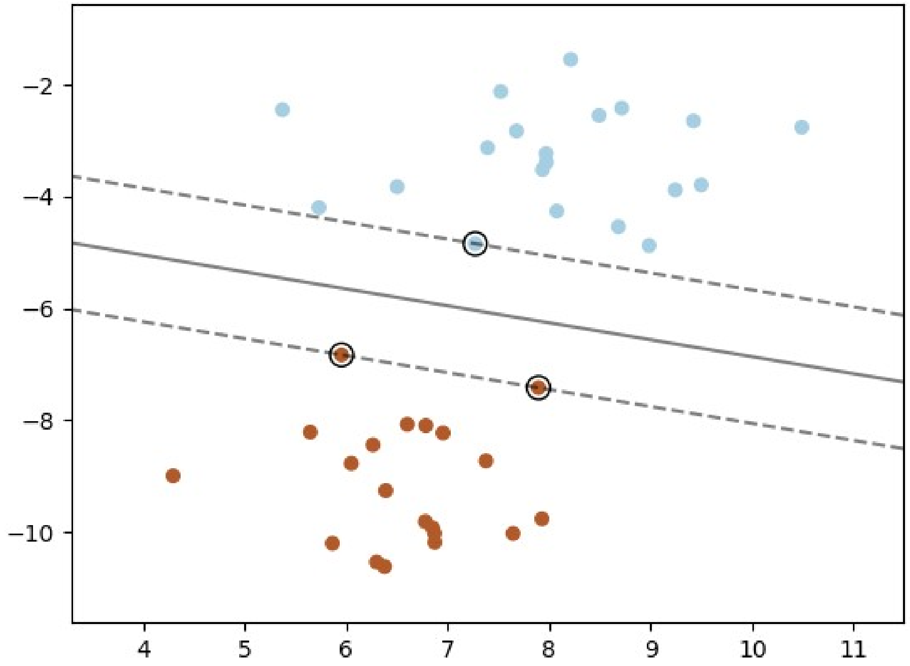


1990s · LeNet

Handwritten document recognition

Late 1990s–early 2000s: declining interest

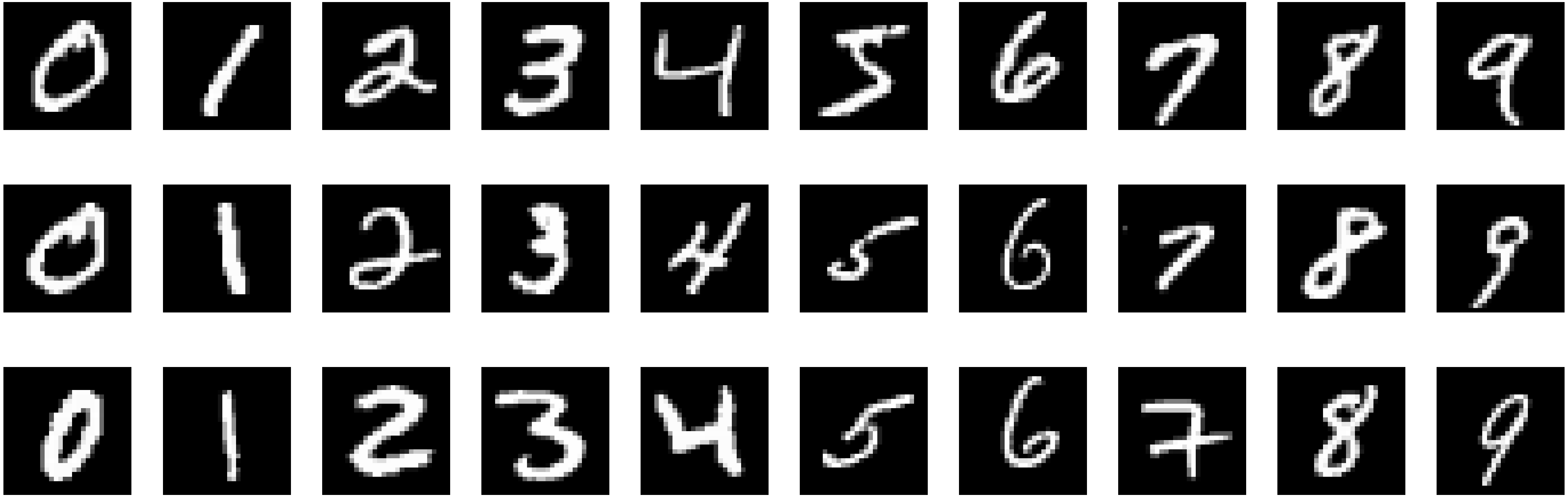
Support vector machines



- Deeper networks were difficult to train reliably
- SVMs offered strong results and well-understood optimization
- Boosting and random forests became popular alternatives

Learning from examples

0 1 2 3 4 5 6 7 8 9



Hand-designed features

Color

Fraction of pixels that are red

Boundary shape

Corners and edge directions

Local appearance

Contrast patterns around lettering

What the measurements miss

Color

Many objects are red.

Boundary

The outline changes with viewpoint.

Lettering

Blur or occlusion can hide the text.

Bag-of-words

The **dog** chased the cat

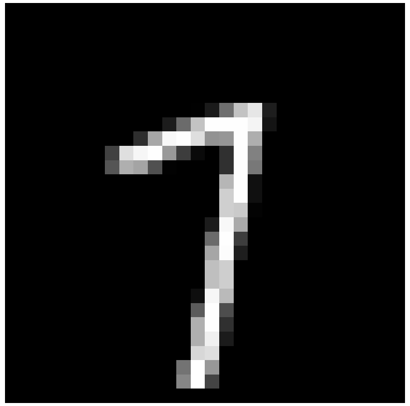
The **cat** chased the dog

Both sentences have the same word counts

the	dog	chased	cat
2	1	1	1

High-dimensional inputs

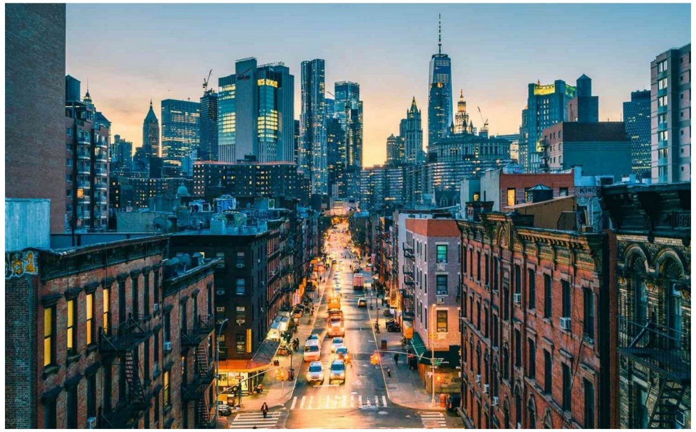
A small digit image



28×28

784 values

A color photograph



$224 \times 224 \times 3$

150,528 values

A DNA sequence

...
AGCTGCGA
GTGGTATT
ACAGGTCT
AGCGCGAG
TTGAAACA
TAGAGGGG
CGCTCGGT
TATCAACA
...

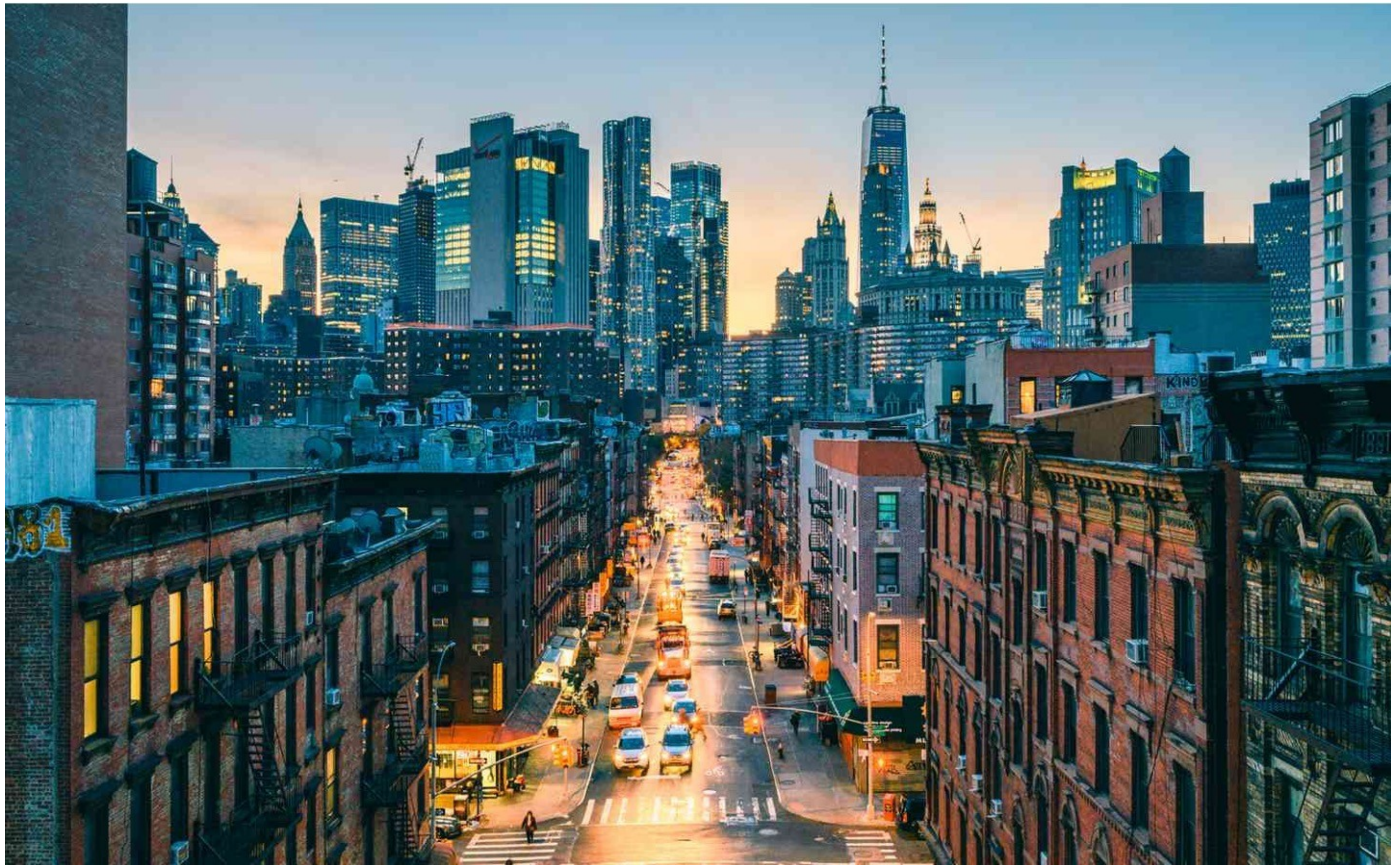
A long DNA input

About 1 million letters

The number of possibilities grows quickly

1 position	10
2 positions	100
10 positions	10,000,000,000

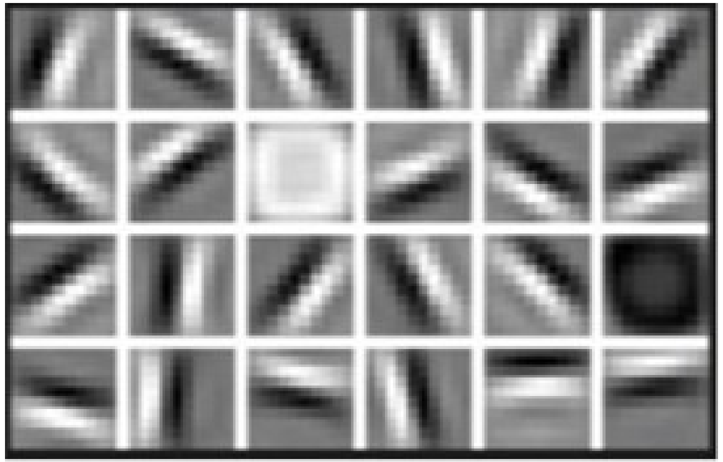
Structure in real data



- Nearby pixels are related
- Shapes and parts recur
- Learned patterns can be reused

Learning features from data

Local patterns



Parts

Larger structures

- Intermediate descriptions can be learned from data
- Several learned stages can combine simpler patterns

2006: renewed interest in deep networks

Deep belief networks

Hinton, Osindero and Teh

Generated samples of handwritten digits

ImageNet

1,000 categories

1.2 million training images

ImageNet classification

Container ship

Leopard

1,000 categories

Top-1

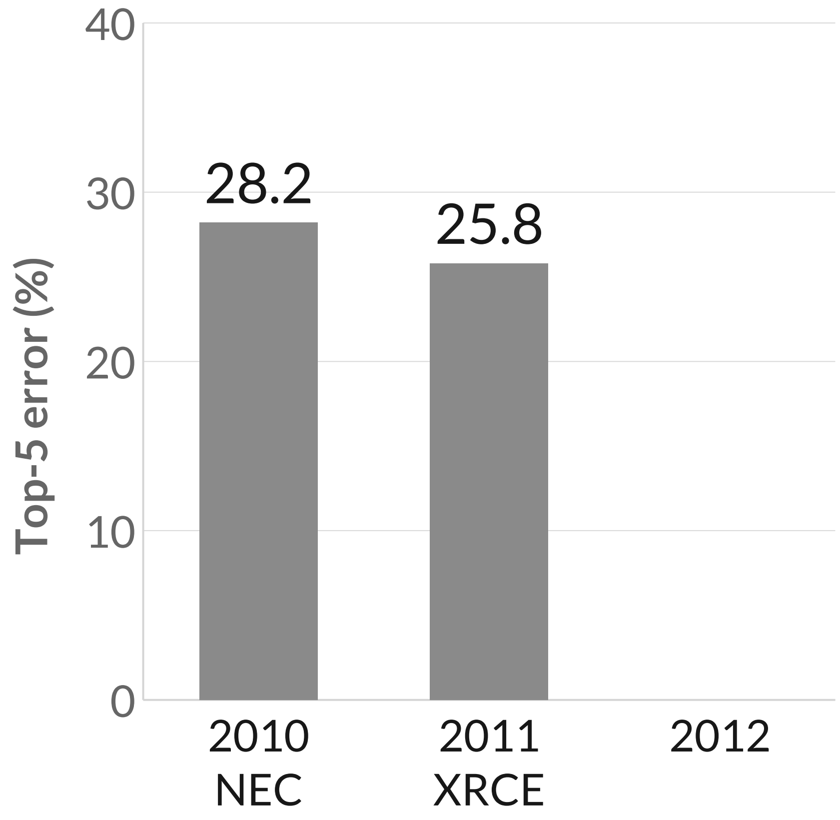
First proposed category

Top-5

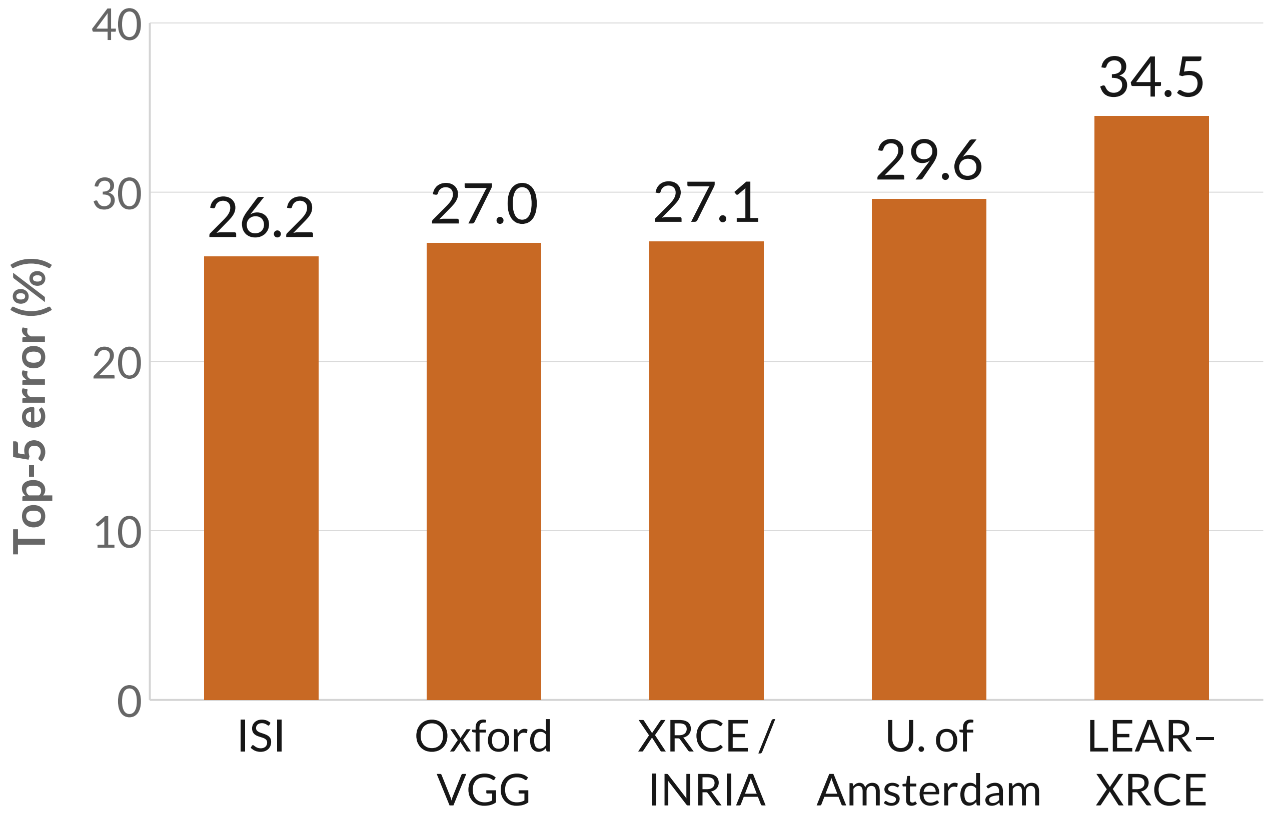
Correct label in
the top five

ILSVRC-2012 Leaderboard

Winning entries



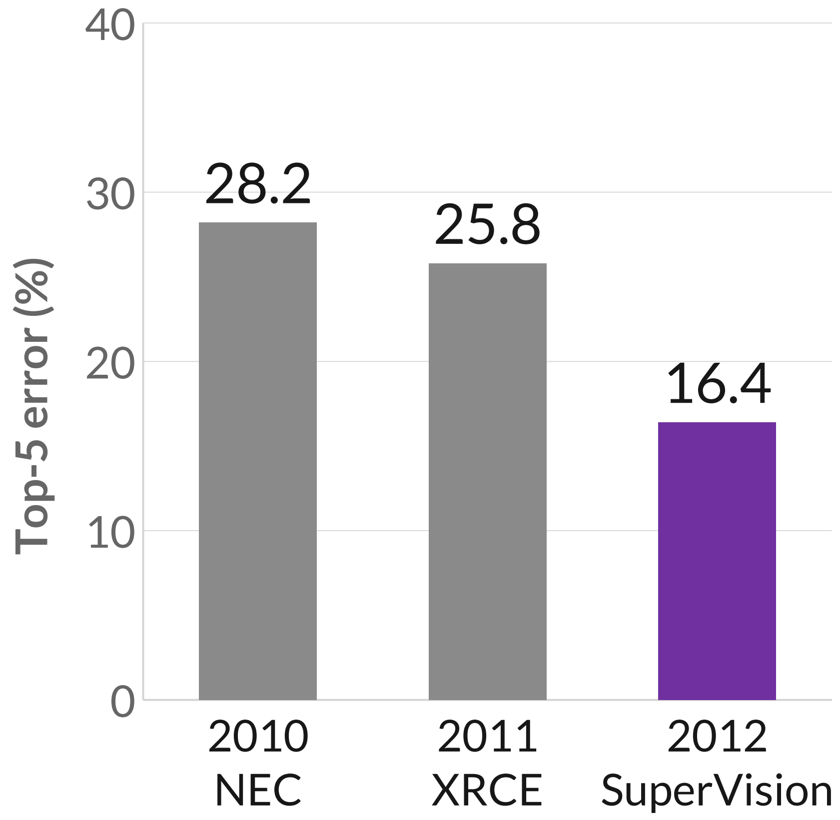
Other 2012 entries



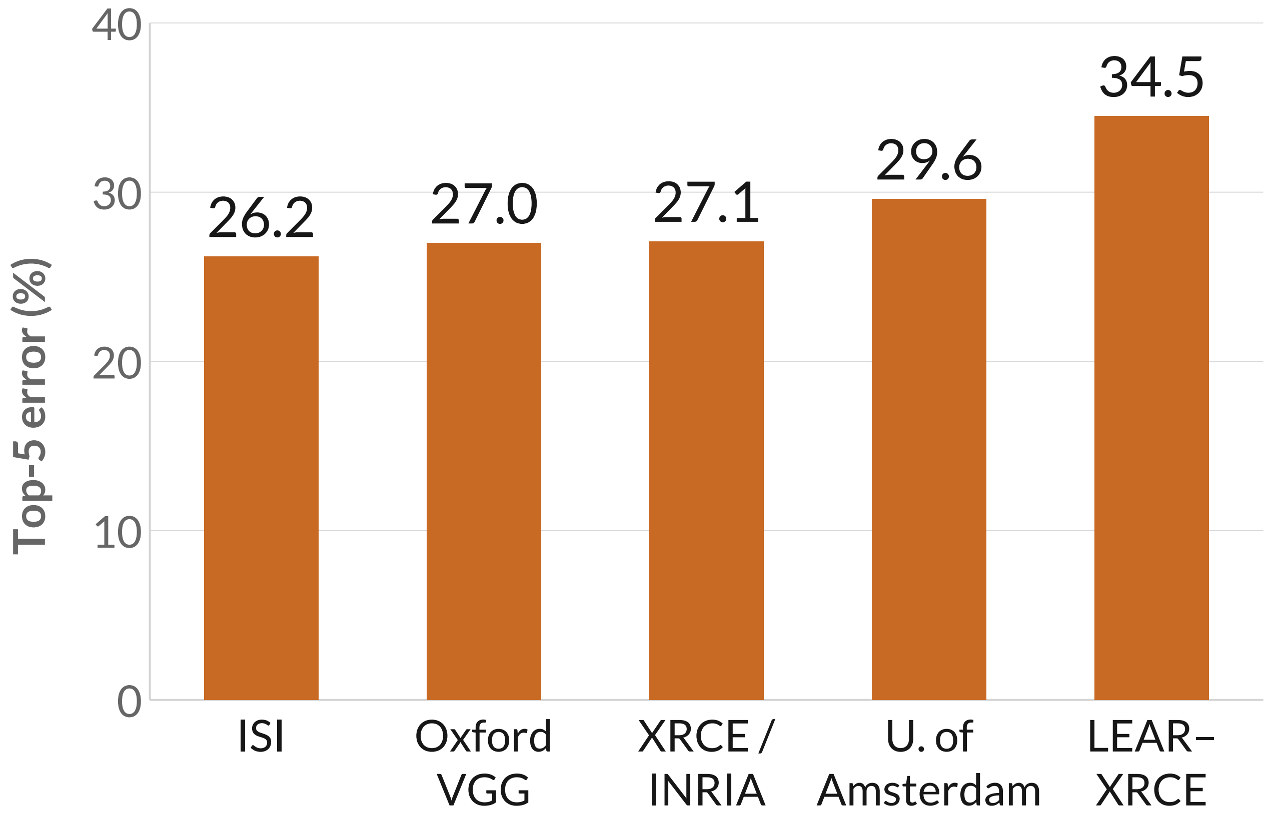
ILSVRC-2012 Leaderboard

Winning entries

SuperVision = AlexNet



Other 2012 entries



AlexNet: making the experiment practical

Alex Krizhevsky

Ilya Sutskever · Geoffrey Hinton

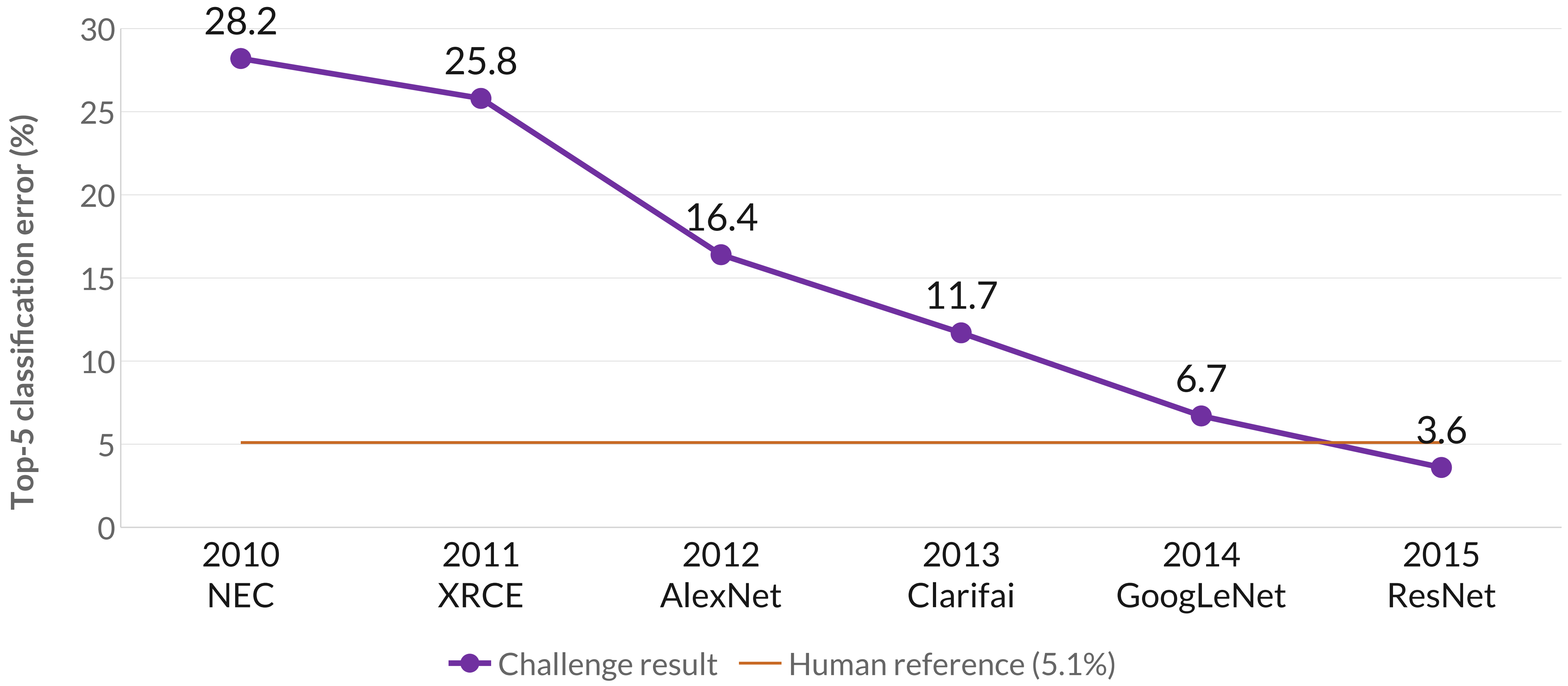
2 GPUs

NVIDIA GTX 580

5–6 days

Training time

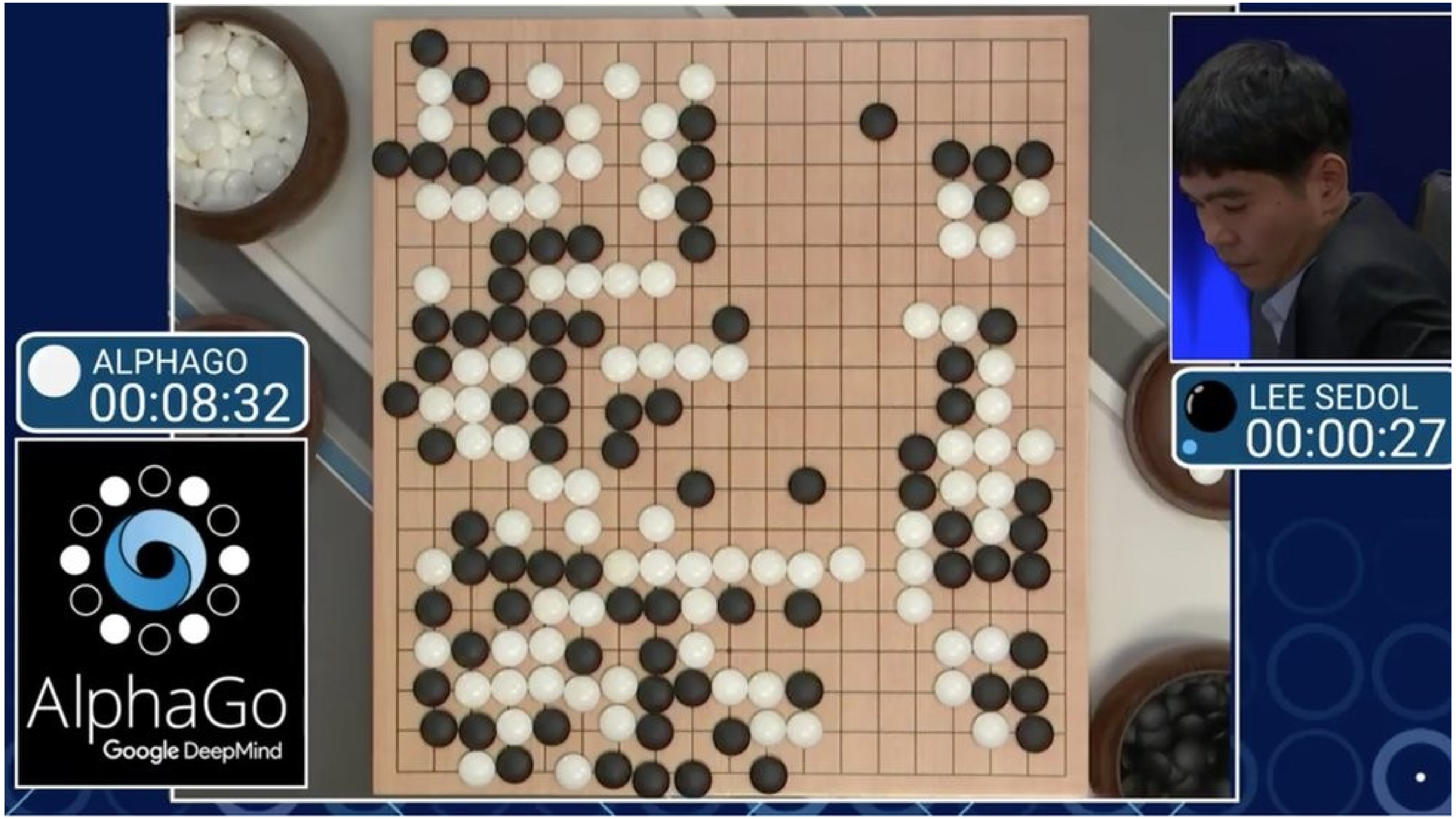
ImageNet recognition, 2010–2015



Deep learning beyond image recognition

2016	**AlphaGo**	Learning and search
2017	**Transformer**	A neural-network design for sequences
2020–21	**AlphaFold 2**	Protein structure prediction
2020s	**Large language models**	A general language interface

AlphaGo



Lee Sedol vs. AlphaGo, 2016

Learning to continue text

I forgot my umbrella.

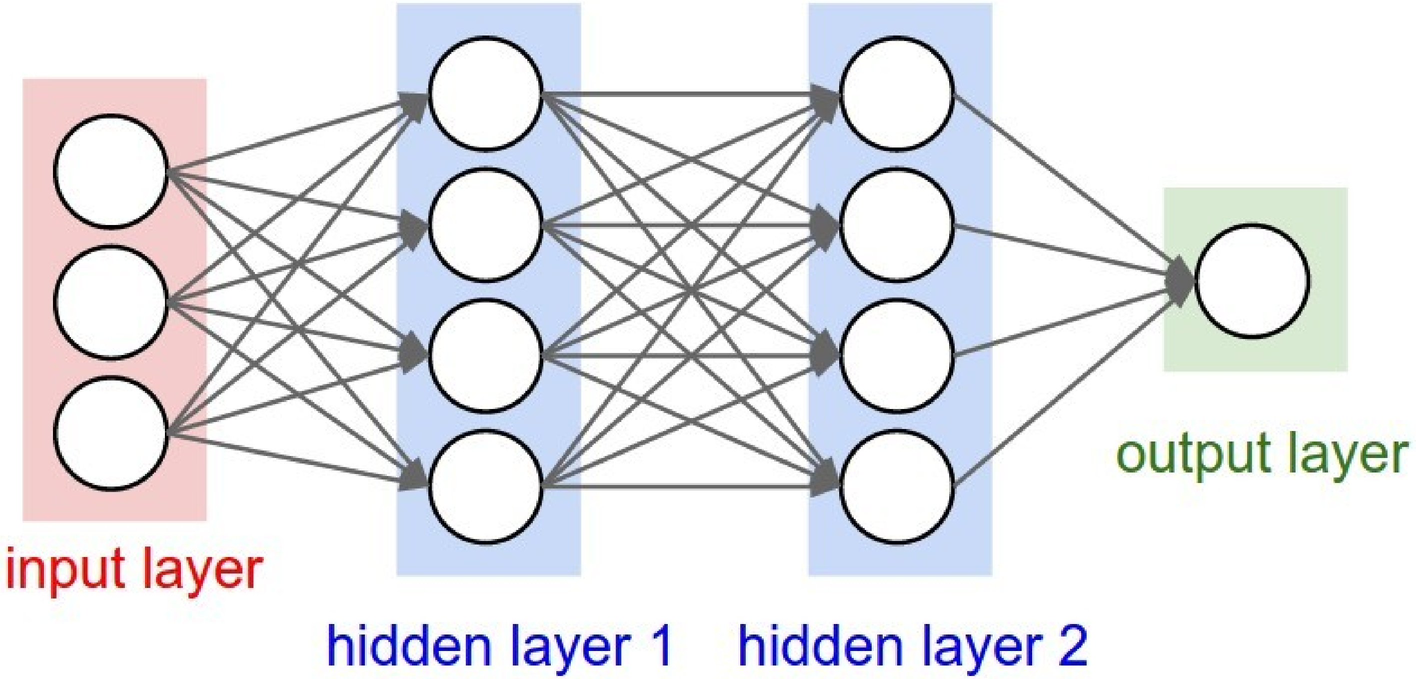
By the time I reached campus, I was …

soaked wet shivering

The data can provide the learning target

Context	Continuation from the original text
Paris is the capital of	**France.**
A triangle has three	**sides.**

Training and using a model



Training

Adjusts parameters using training data

Inference

Applies the learned model to new inputs

One model, different instructions

Translate into Chinese

The train leaves at eight.

火车八点出发。

Explain in one sentence

What is overfitting?

A model fits its training examples but performs poorly on new ones.

Language models and assistants

Answer a question

“How could I plot these measurements?”

Carry out a task

“Open the data, make the plot, and save it.”

Recent developments (2026)

January

AlphaGenome

DNA sequence → molecular predictions

April

π0.7

More general robot skills and instructions

September

GPT-6 Astra

Language models that work with software

Images and language

A white teddy bear sitting in the grass

A man riding a wave on a surfboard

Text to image

DALL·E, 2021: an early text-to-image example

Generating and editing images

Video generation

Seedance 2.0

2026

10-second clip

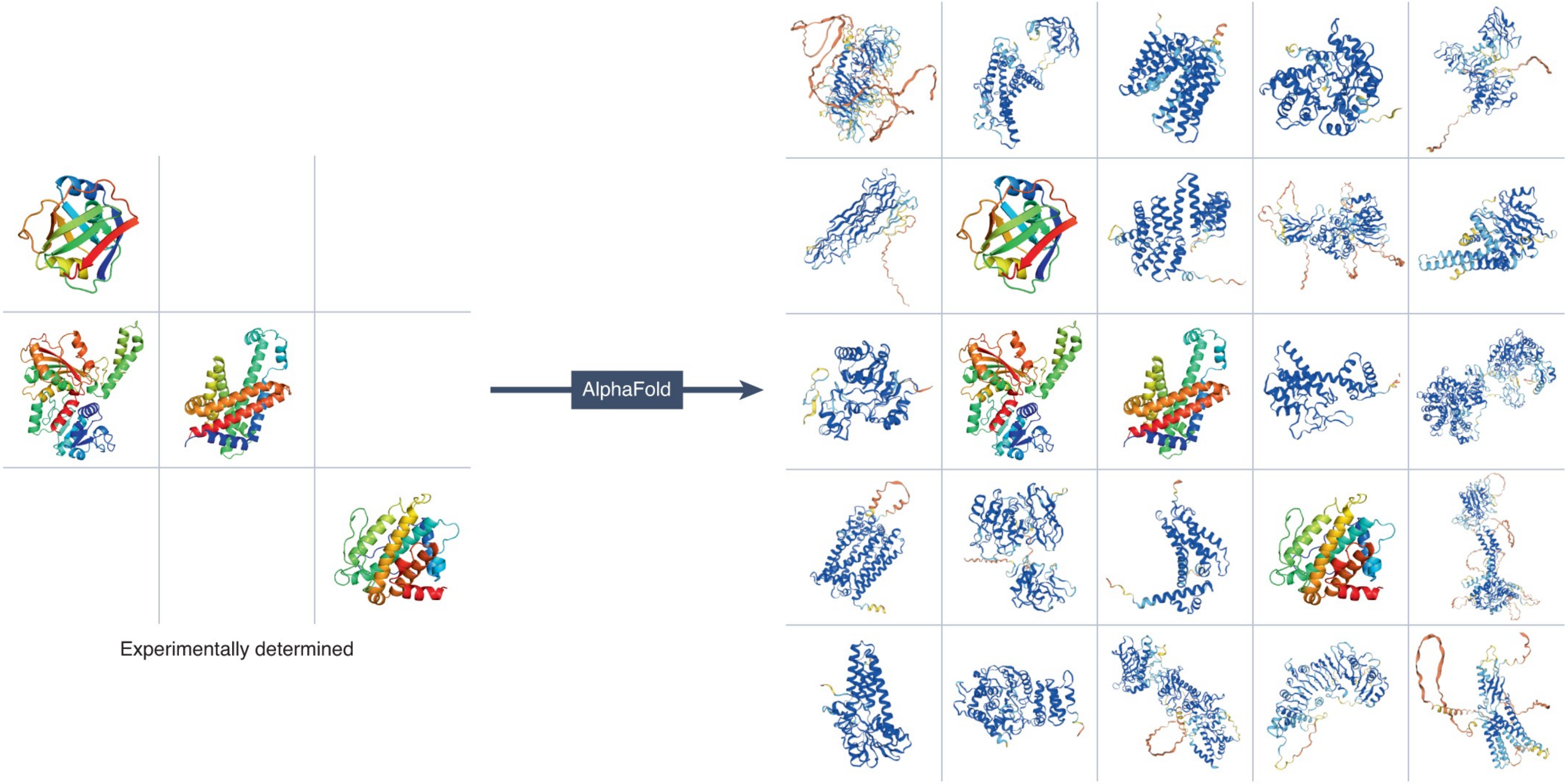
3D modeling with GPT-6 Astra

Editable 3D scene

Built and revised using Blender

30-second camera tour

Predicting protein structures



Experimentally determined

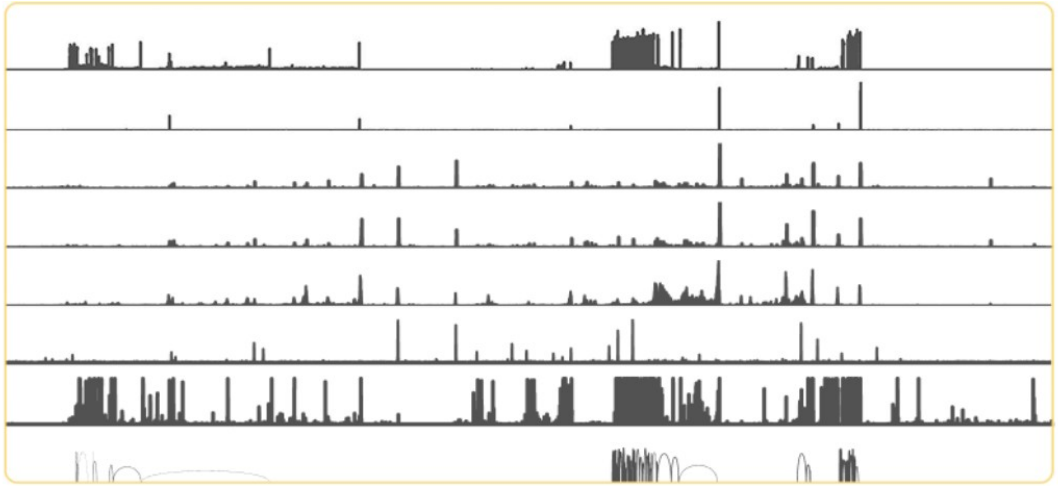
Expanded coverage with structure prediction

Predicting molecular activity from DNA

DNA sequence

...
AGCTGCGA
GTGGTATT
ACAGGTCT
AGCGCGAG
TTGAAACA
TAGAGGGG
CGCTCGGT
TATCAACA
...

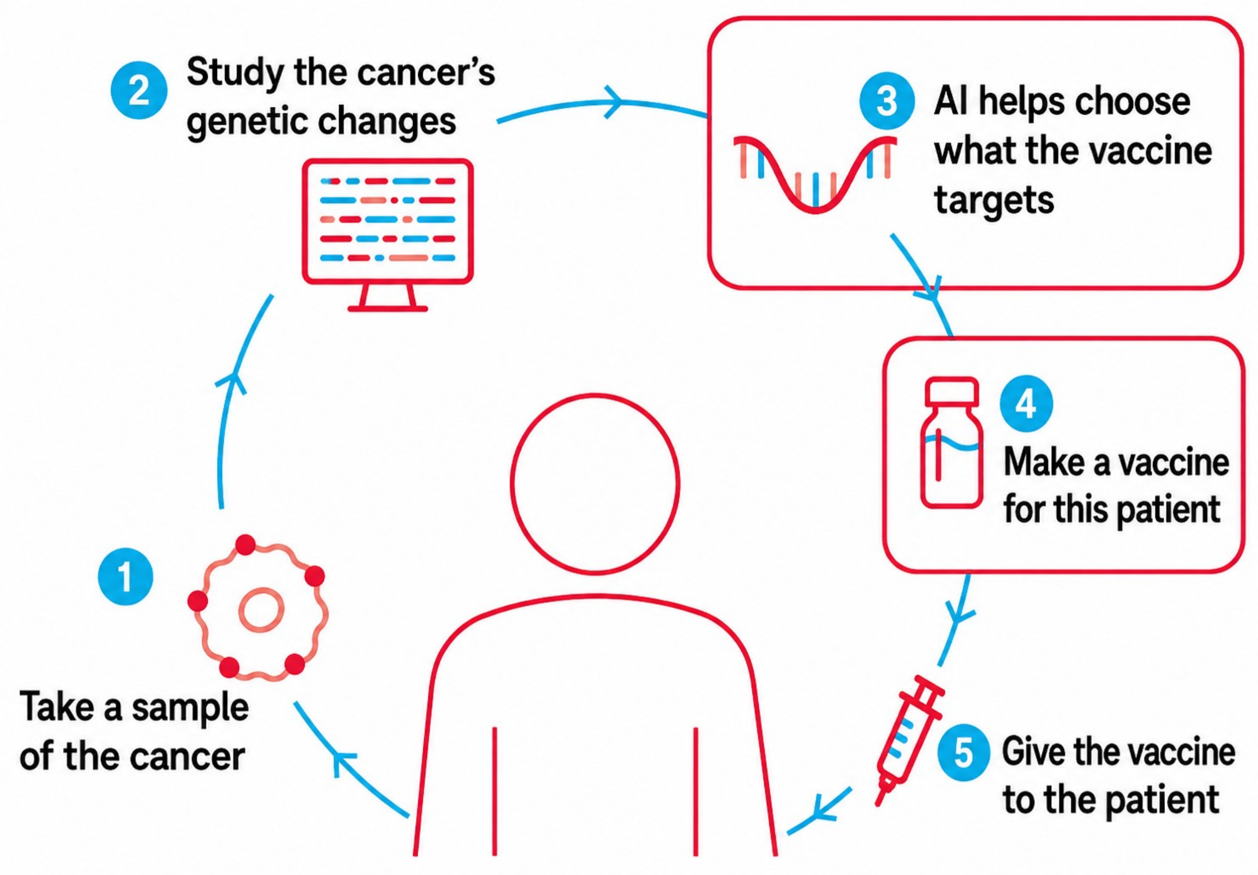
Predicted molecular activity



AlphaGenome · 2026

AI in personalized cancer vaccines

Moderna + Merck · Experimental treatment



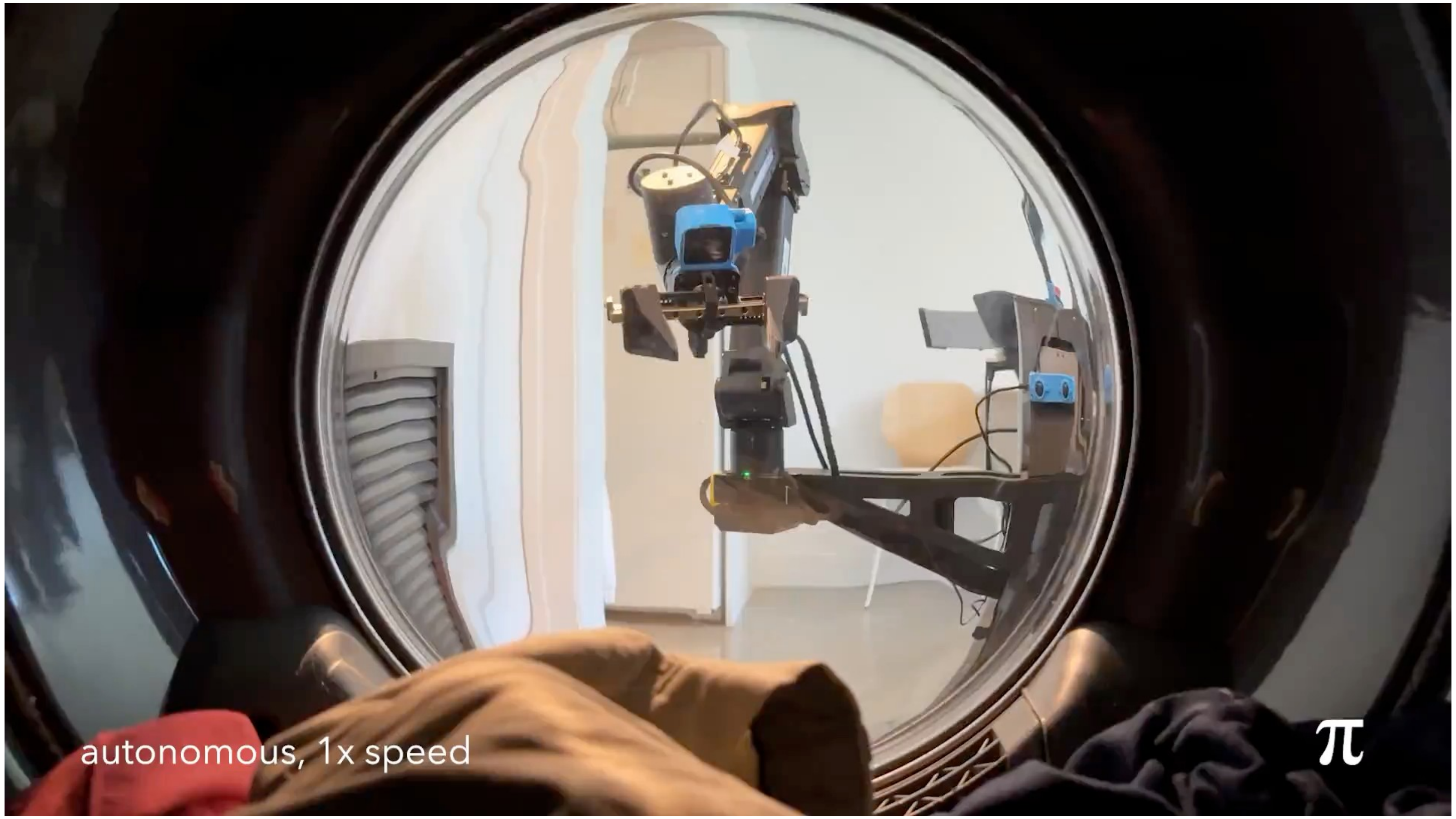
A vaccine designed for each patient

AI helps choose what the immune system should learn to recognize.

Skin-cancer trial · 2026

Adding the vaccine to standard treatment helped patients stay cancer-free for longer.

Robot learning



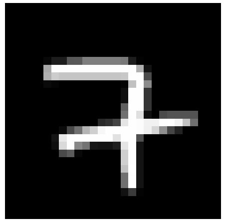
Physical Intelligence π0 demonstration, 2024

Robot learning in 2026

π0.7 · April 2026

Learning signals

Supervised



Label: 7

A supplied target

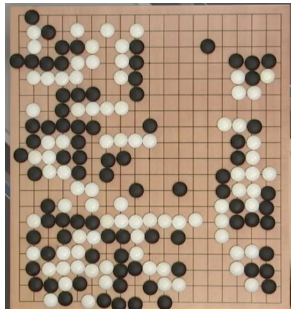
Self-supervised

The cup is on the …

table

The next word comes from the original text

Reinforcement



Actions change the board

Feedback: win or lose

Deep learning is science and engineering

Data

Text, video and task experience

Collection, filtering, duplicate removal

Models

Network design and training

Using context across text or time

Training systems

Computation at a practical scale

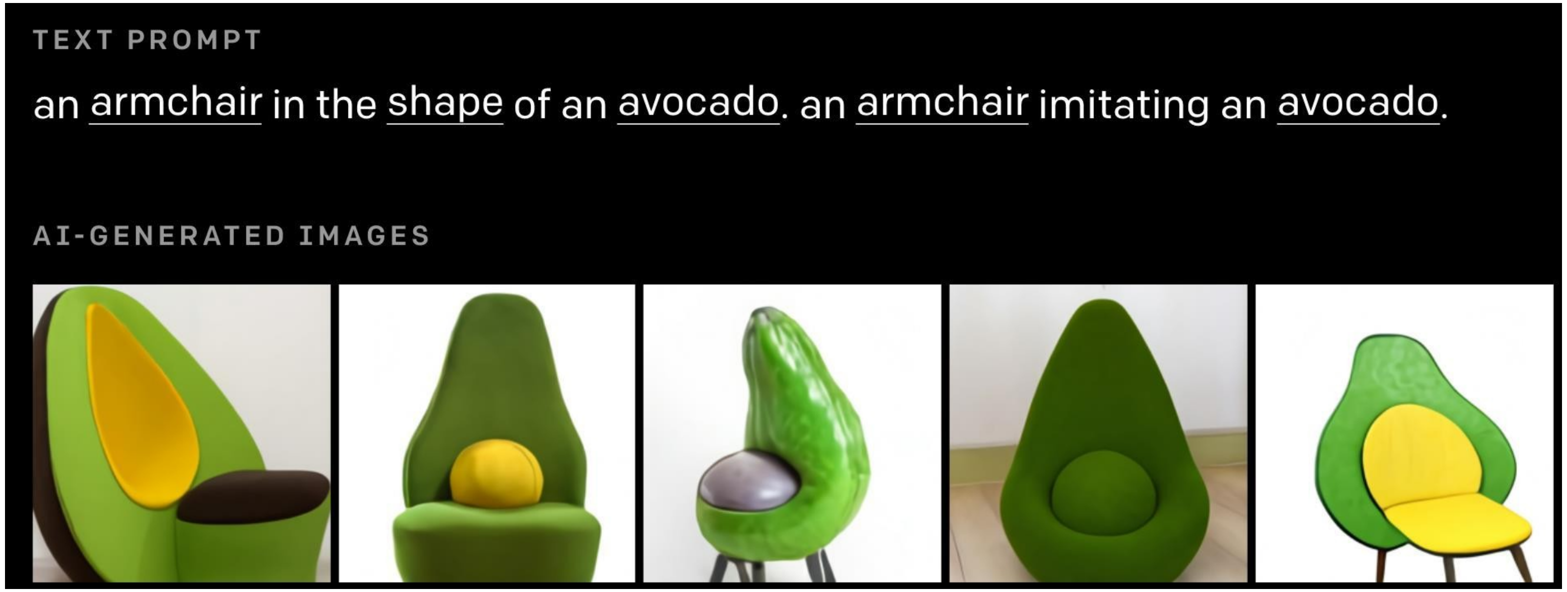
GPU memory and parallel execution

Evaluation

Performance on unfamiliar examples

Correctness, consistency and cost

Evaluating image generation



- Does the image follow the prompt?
- Can several different outputs be valid?

When the input data changes

Handwritten digits

Street numbers

Memorization and generalization

A familiar fact

$$7 \times 8 = 56$$

Recall the answer

A new calculation

$$17 \times 8 =?$$

$$(10 + 7) \times 8$$

$$= 80 + 56 = 136$$

Learning from experience

David Hume · 1748

Past observations

The sun has risen on every day we observed.

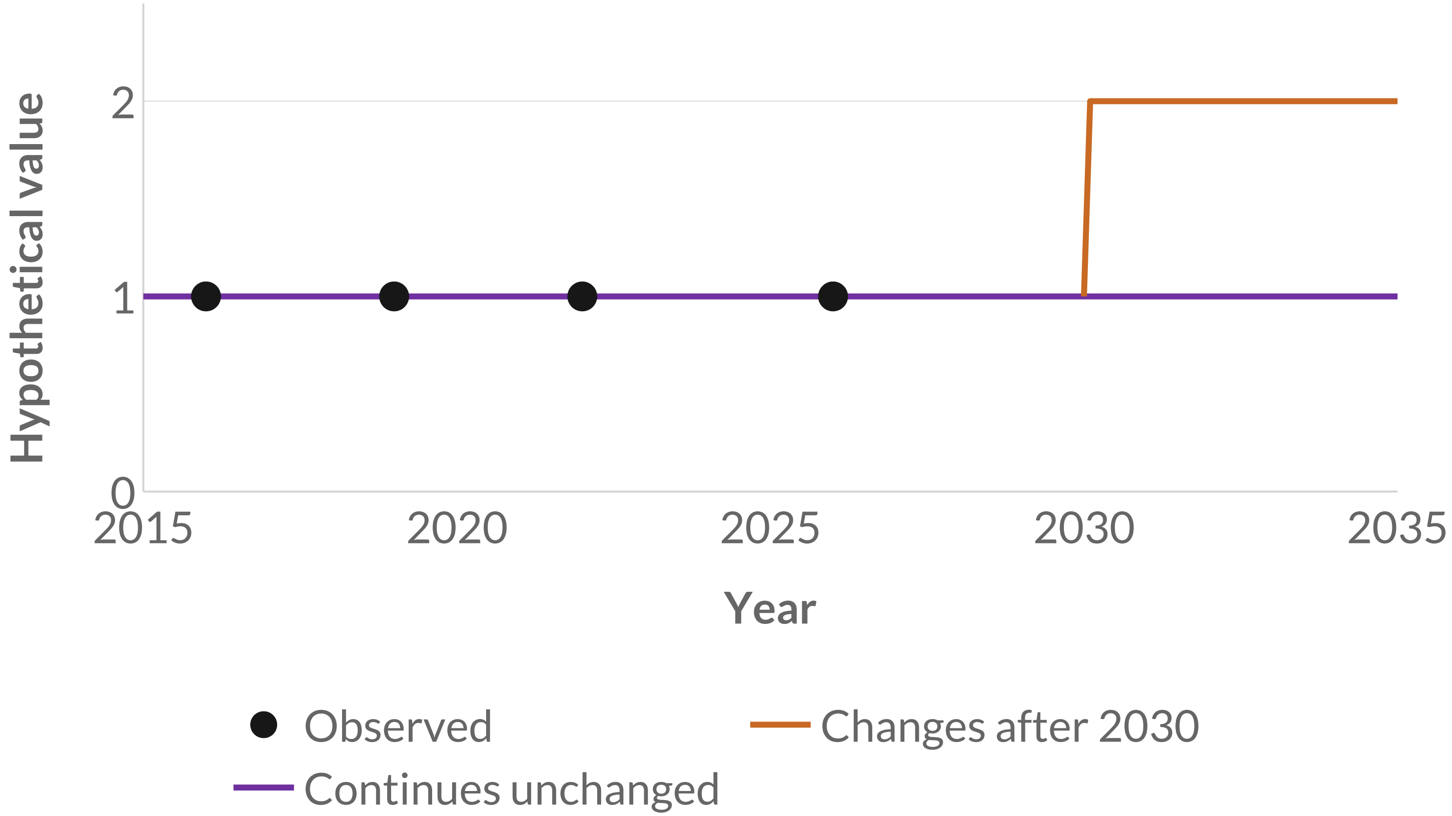
An unobserved case

The sun will rise tomorrow.

Strong support is different from a logical guarantee.

Different rules fit the same observations

Inspired by Nelson Goodman · 1955



Both rules fit all observed points.

The predictions differ after 2030.

Prior assumptions in learning

$$2, \quad 4, \quad 6, \quad 8, \quad \ldots$$

Repeated rule

Add 2 each time

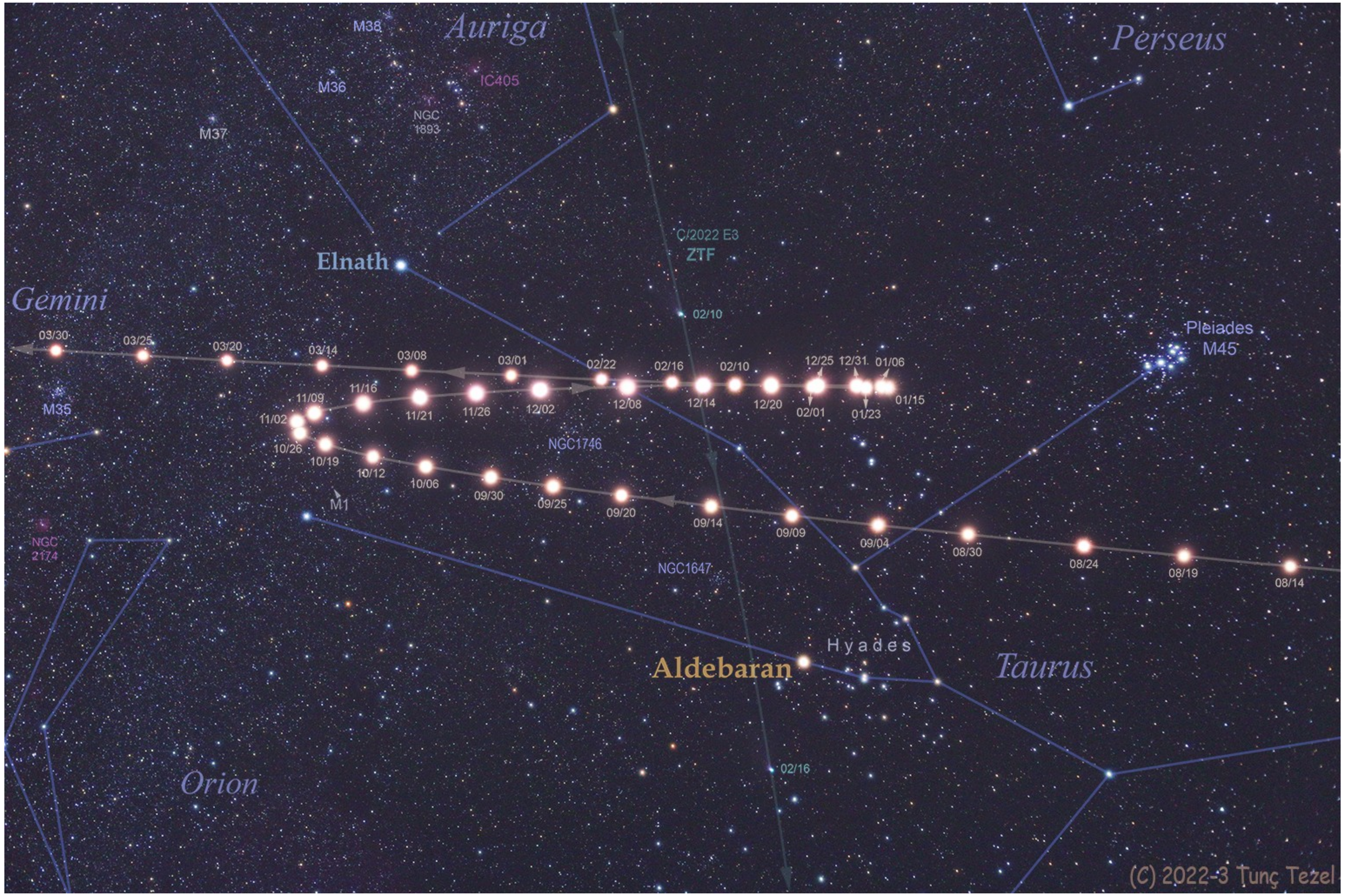
$$10$$

Rule with an exception

Add 2 through term 4,
then add 100

$$108$$

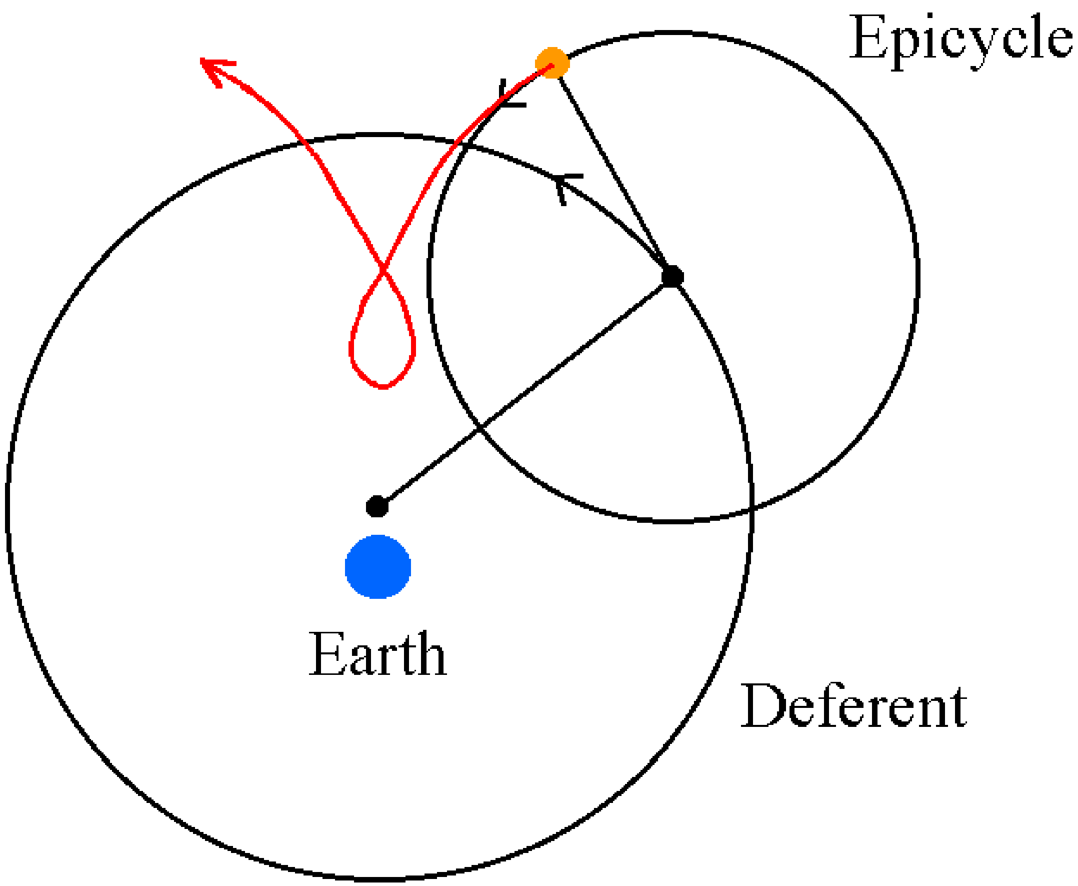
Mars against the background stars



Mars's position
at different dates

Track Mars
against the stars.

The Earth-centered model



- The planet follows the small circle.
- Its center circles Earth.
- Together, the motions can produce a backward loop.
- More corrections make the model more complex.

Occam's razor

Prefer a simpler explanation when it accounts for the evidence comparably well.

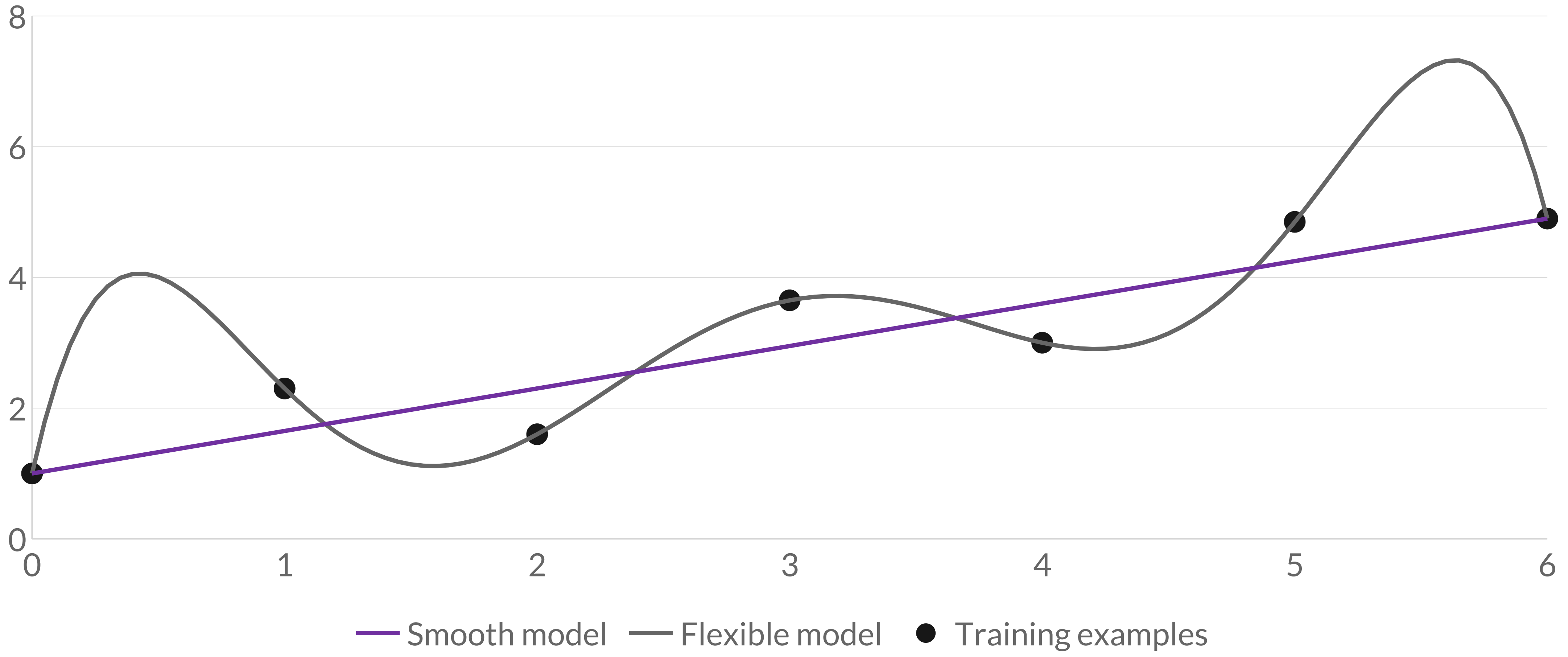
Repeated rule

Add 2 at every step

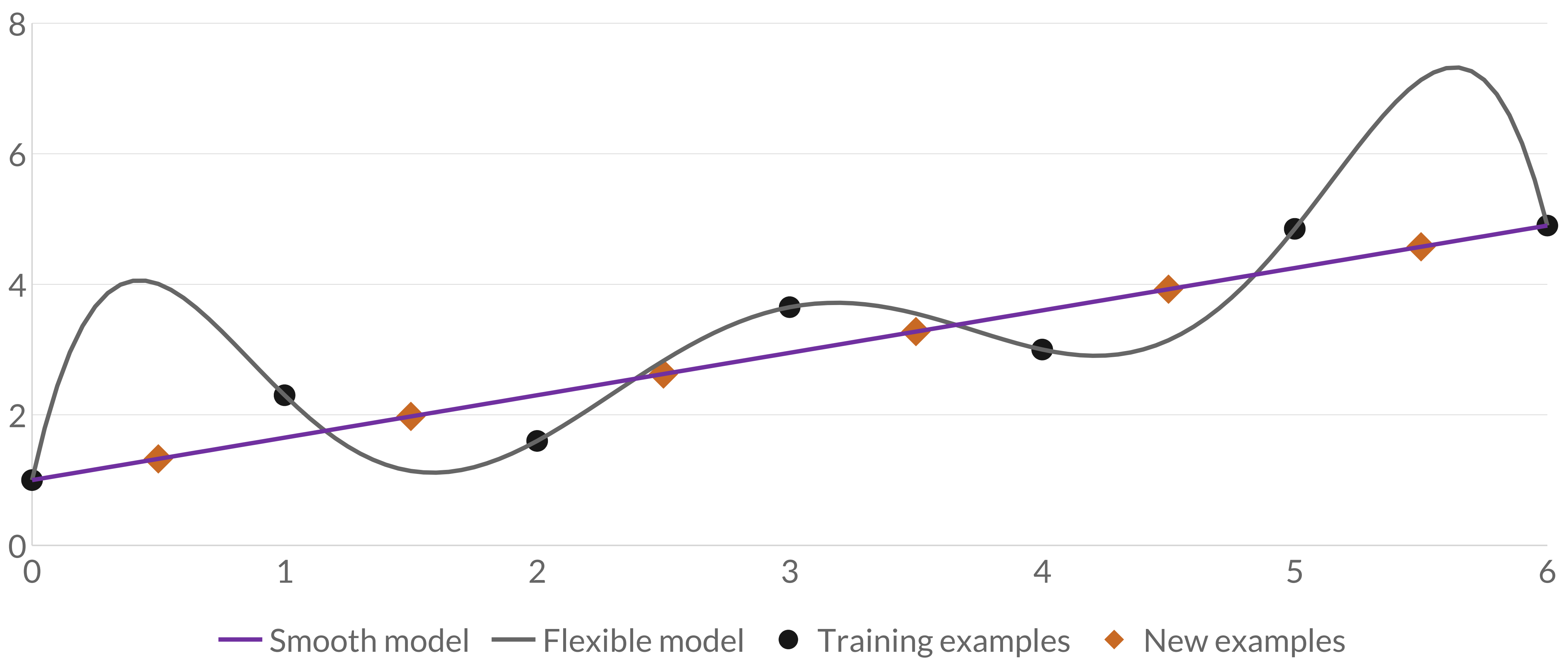
Extra condition

Change the rule just after the observed examples

Fitting the observed examples

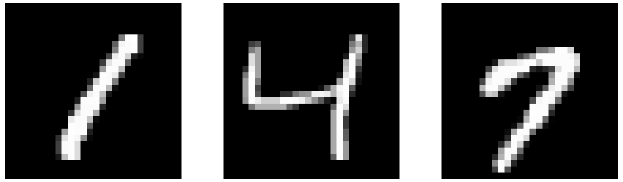


Predicting new examples



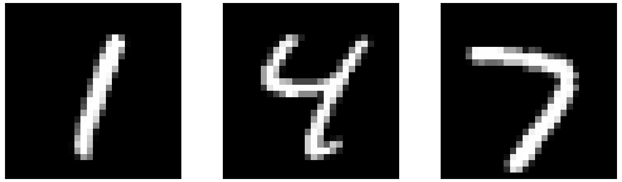
Training, validation and test data

Training



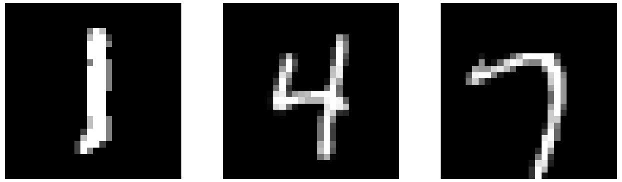
Fit the model

Validation



Choose settings

Test



Final evaluation

Studying a deep learning system

- **Model:** The computation and its assumptions
- **Learning:** How data changes the parameters
- **Evaluation:** What works on new examples

What you will learn

- **Understand:** Follow network computation and learning
- **Train:** Build models and diagnose their behavior
- **Choose:** Connect model designs to data and tasks
- **Apply:** Develop and evaluate a small application

What this course is not about

- A full PyTorch programming tutorial
- A catalogue of recipes for every dataset
- A survey of every new model or product

The first half of the course

- **Weeks 1–2:** Neural networks and learning
- **Weeks 3–5:** Vision models and visual tasks
- **Weeks 5–6:** Sequences, attention and Transformers
- **Week 7:** Review and midterm

The second half of the course

- **Weeks 8–9:** Self-supervision and generative models
- **Weeks 10–11:** Reinforcement learning and foundation models
- **Weeks 12–13:** 3D, robotics, graphs and AI for science
- **Week 14:** Applications, project and review

Prerequisites

- **Machine learning:** Models, supervision and generalization
- **Linear algebra:** Vectors and matrix multiplication
- **Calculus:** Derivatives and the chain rule
- **Probability:** Random variables and expectations
- **Programming:** Python and numerical computing

Studying deep learning

- Understand the idea
- Follow a small example
- Implement and experiment

Reading

No required textbook

Deep Learning

Ian Goodfellow · Yoshua Bengio · Aaron Courville

MIT Press

Deep Learning: Foundations and Concepts

Christopher M. Bishop · Hugh Bishop

Springer

Lectures and instructor

Tuesday & Thursday

10:30–12:00

J_E205 教学楼 E205

Zhen Liu

zhenliu@cuhk.edu.cn

Instructor office hours

Tuesdays, 09:00–10:00, every other week from 15 September
Location to be announced
By request. Email at least one day in advance.

Teaching assistants and office hours

Wenqian Zhang 张文谦

Lead TA · 225040483@link.cuhk.edu.cn

Tue 17:00–18:00 · Thu 14:30–15:30

Zhixin 408
知新楼 408

Ruixiang Wang 王瑞翔

TA · 225040514@link.cuhk.edu.cn

Sat 17:00–18:00 · Sun 17:00–18:00

Teaching Complex B 609
教学综合楼 609

Yifan Peng 彭一凡

TA · 225040521@link.cuhk.edu.cn

Mon 19:00–20:00 · Wed 19:00–20:00

Zhixin 429
知新楼 429

Emailing the teaching team

Start every subject line with

[DDA4220]

[DDA4220] Question about Assignment 1

[DDA4220] Request for carrying cats to the class

Assessment

70%

Coursework

45% 3 homework assignments

25% Final project

30%

Midterm

Tentative date

Thu, 22 October 2026

Coursework and collaboration

- Start early enough to investigate the result
- Follow each assignment’s collaboration rules
- Acknowledge permitted external material
- Submit work that you understand and wrote

Assignment bonus challenges

- Each of the three assignments includes a bonus challenge
- Credit for performance, ideas and leaderboard results
- Submit reproducible work and explain your approach
- The core assignment can earn full marks on its own

The course project

- Teams of 3–4 students; find collaborators early
- Choose any topic related to deep learning
- Investigate a substantive problem beyond an assignment
- University cluster access for queued GPU jobs

Project directions

Applications

Develop and evaluate a system for a meaningful task.

Methods and systems

Explore changes to a model, training method or implementation.

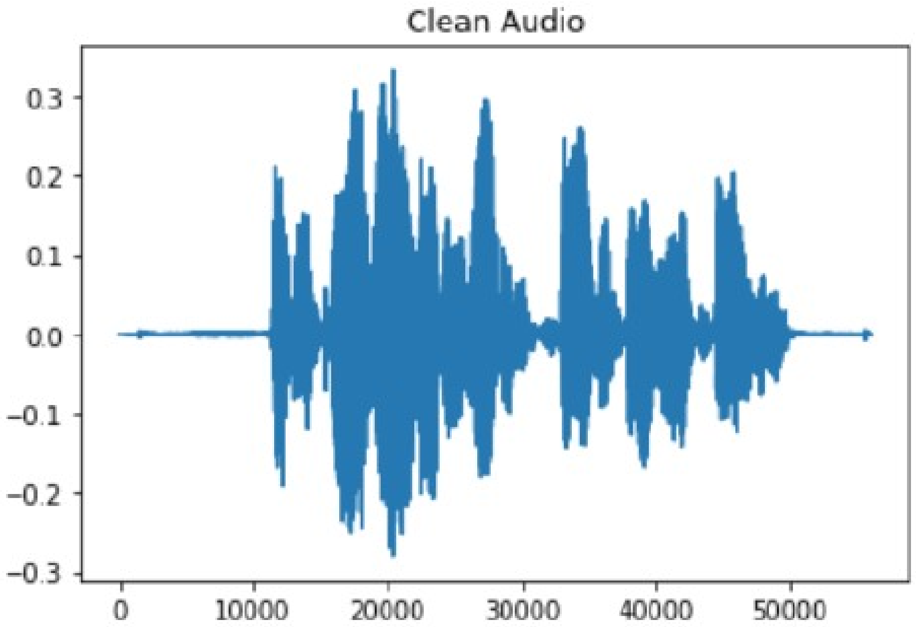
Empirical studies

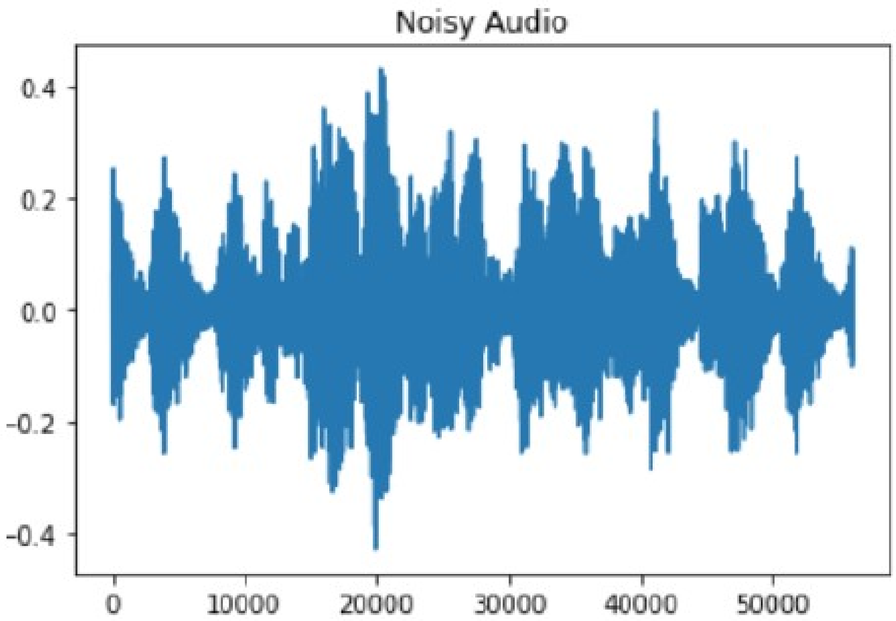
Investigate model behaviour through controlled experiments.

Running existing code alone is not sufficient.

Example project: speech enhancement

A student project from Stanford CS230





- Prepare speech recordings with realistic background noise
- Train a model and investigate its design choices
- Evaluate noise removal and preservation of the speaker’s voice

Project assessment

- Substantial work and clear individual contributions
- Relevant prior work and your project’s position
- A worthwhile, well-defined goal or question
- Sound experiments and evaluation
- A clear, carefully formatted report

Project milestones

Shares of the project assessment

20%	30%	50%
Proposal	Progress report	Final presentation and report

Conference-paper report format · Dates to be announced

Working as a team

- Find collaborators early; we can help form teams
- Agree on responsibilities and review the work together
- Keep communication and code-contribution records
- Contact the TAs early about collaboration problems

AI use in coursework

For understanding

Explain a concept.
Clarify something from a lecture.

No AI-written assignments

No generated answers,
derivations, code or reports.

Next lecture: neural networks

- Multilayer perceptrons and activation functions
- Expressive capacity of neural networks
- Forward computation and loss functions
- Backpropagation and parameter updates