Introduction to Artificial Intelligence

Lecture 0: Artificial Intelligence

Prof. Gilles Louppe
g.louppe@uliege.be

LIÈGE université

ChatGPT PLUS

Help me pick
a gift for my dad who loves fishing

Brainstorm edge cases
for a function with birthdate as input, horoscope as ou...

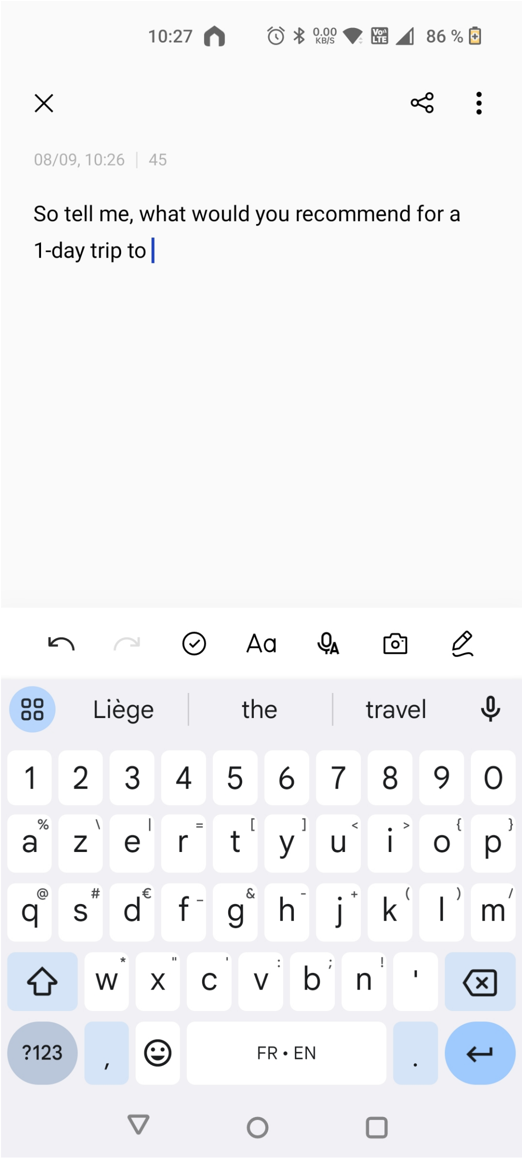
Make up a story
about Sharky, a tooth-brushing shark superhero

Create a personal webpage for me
after asking me three questions

Send a message

One simple idea:

Guess the next word



In the 1960s, Armstrong _____

In the 1960s, Armstrong performed ___

In the 1960s, Armstrong performed a moonwalk ___

In the 1960s, Armstrong performed a moonwalk on the ___

In the 1960s, Armstrong performed a moonwalk on the lunar ____

In the 1960s, Armstrong performed a moonwalk on the lunar surface and said ___

This explains why large language models ...

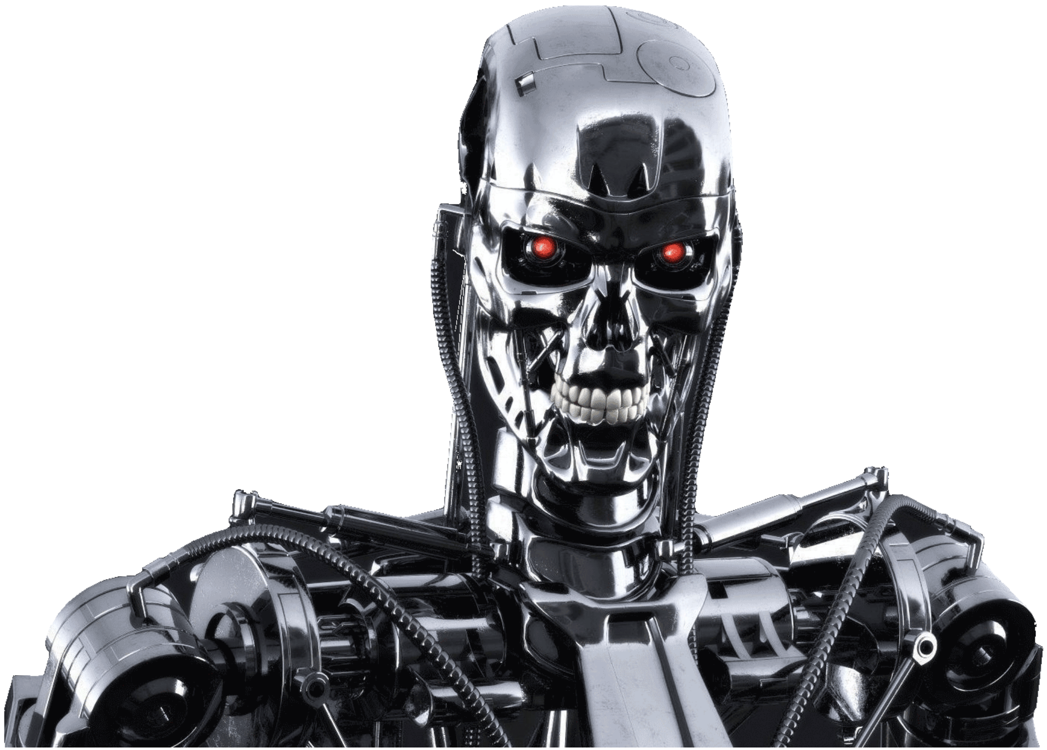
- invent things and cannot cite sources;
- never produce the same answers;
- cannot count, compute, or reason*;
- can hardly correct their own mistakes once they have been made.

—

*: At least not with a vanilla transformer and a greedy decoding strategy.

Not just text, but also images and sounds.

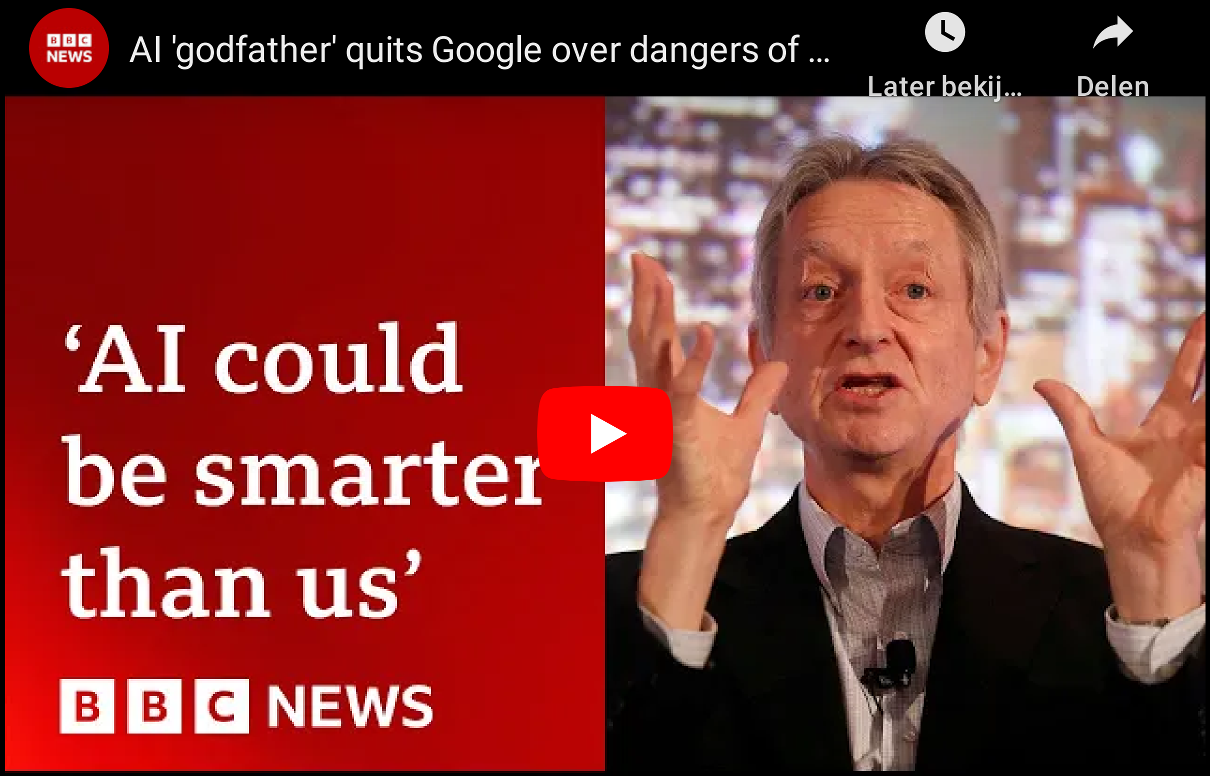
Artificial Intelligence



"With artificial intelligence we are summoning the demon" -- Elon Musk, 2014.

"We're really closer to a smart washing machine than Terminator" -- Fei-Fei Li, Director of Stanford AI Lab, 2017.

Yann LeCun, 2018.



Geoffrey Hinton, 2023.

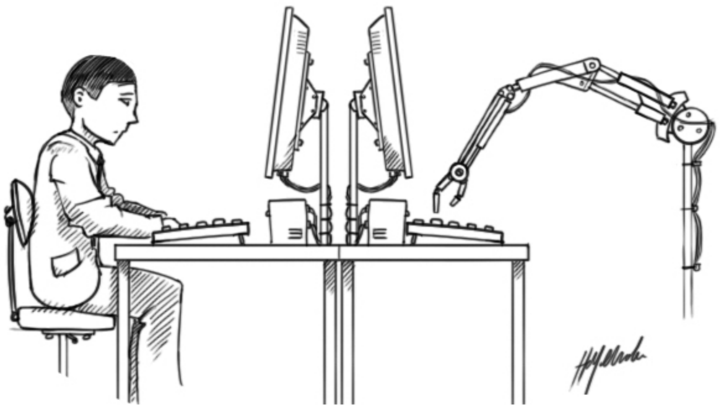
Yann LeCun, 2023.

A definition of AI?

"Artificial intelligence is the science of making machines do things that would require intelligence if done by men." -- Marvin Minsky, 1968.

The Turing test

A computer passes the Turing test (aka the Imitation Game) if a human operator, after posing some written questions, cannot tell whether the written responses come from a person or from a computer.



Can machines think?
(Alan Turing, 1950)

An agent would not pass the Turing test without the following requirements:

- natural language processing
- knowledge representation
- automated reasoning
- machine learning
- computer vision (total Turing test)
- robotics (total Turing test)

Despite being proposed almost 70 years ago, the Turing test is still relevant today.

The Turing test tends to focus on human-like errors, linguistic tricks, etc.

However, it seems more important to study the principles underlying intelligence than to replicate an exemplar.

Aeronautics is not defined as the field of making machines
that fly so exactly like pigeons that they can fool even other pigeons.

A modern definition of AI

An ‘AI system’ is a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. -- European AI Act, Article 3, 2024.

—
Regulation (EU) 2024/1689.

A short history of AI

1940-1950: Early days

- 1943: McCulloch and Pitts: Boolean circuit model of the brain.
- 1950: Turing's "Computing machinery and intelligence".

1950-1970: Excitement and expectations

- 1950s: Early AI programs, including Samuel's checkers program, Newell and Simon's Logic Theorist and Gelernter's Geometry Engine.
- 1956: Dartmouth meeting: "Artificial Intelligence" adopted.
- 1958: Rosenblatt invents the perceptron.
- 1965: Robinson's complete algorithm for logical reasoning.
- 1966-1974: AI discovers computational complexity.

The Darthmouth workshop (1956)

The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described *that a machine can be made to simulate it.*


The Thinking Machine (Artificial Intelligenc...
Later bekij...
Delen

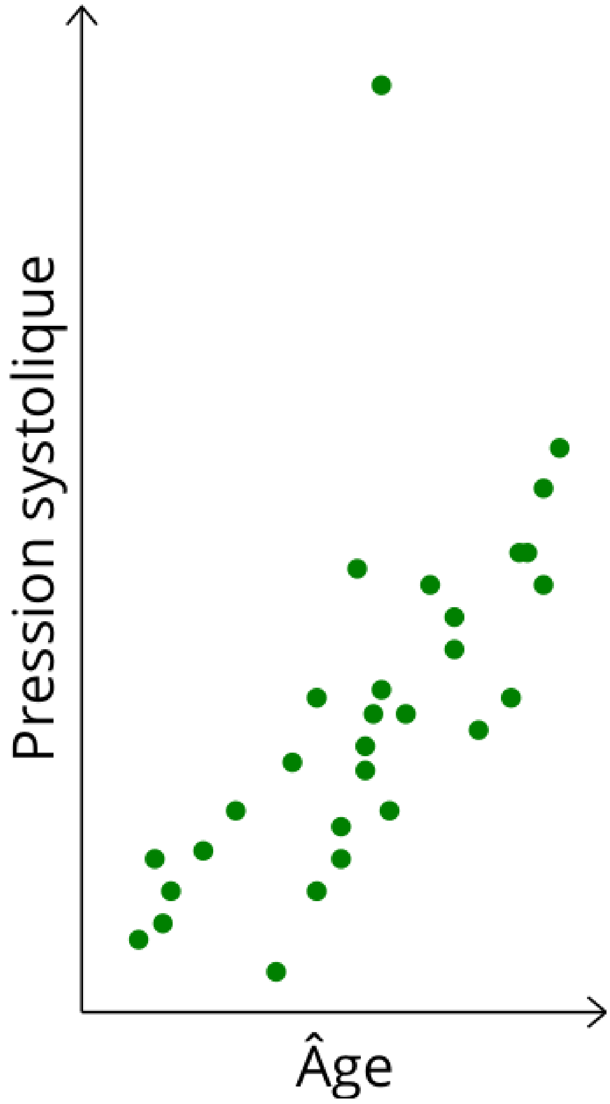
1970-1990: Knowledge-based approaches

- 1969: Neural network research almost disappears after Minsky and Papert's book (1st AI winter).
- 1969-1979: Early development of knowledge-based systems.
- 1980-1988: Expert systems industrial boom.
- 1988-1993: Expert systems industry busts (2nd AI winter).

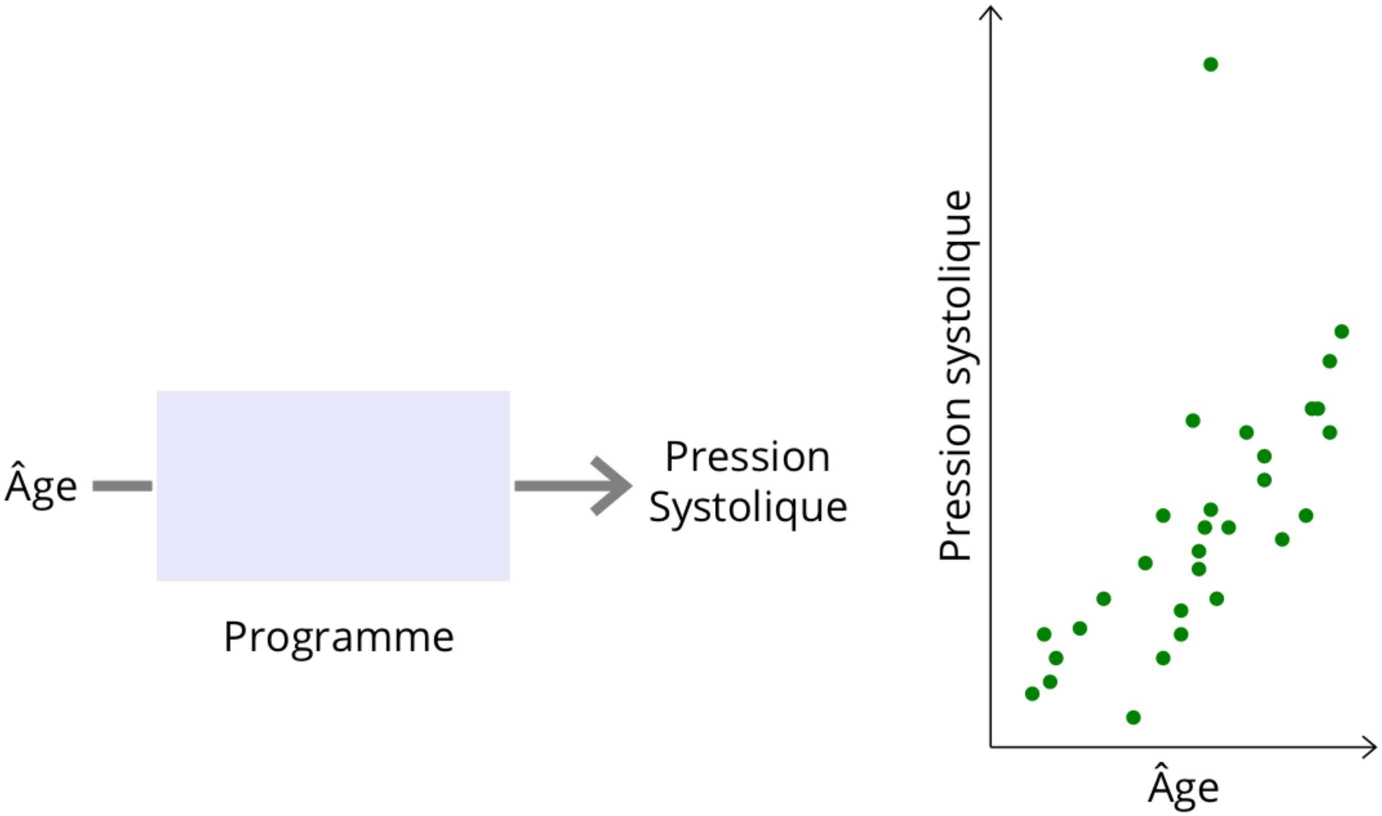
1990-Present: Statistical approaches

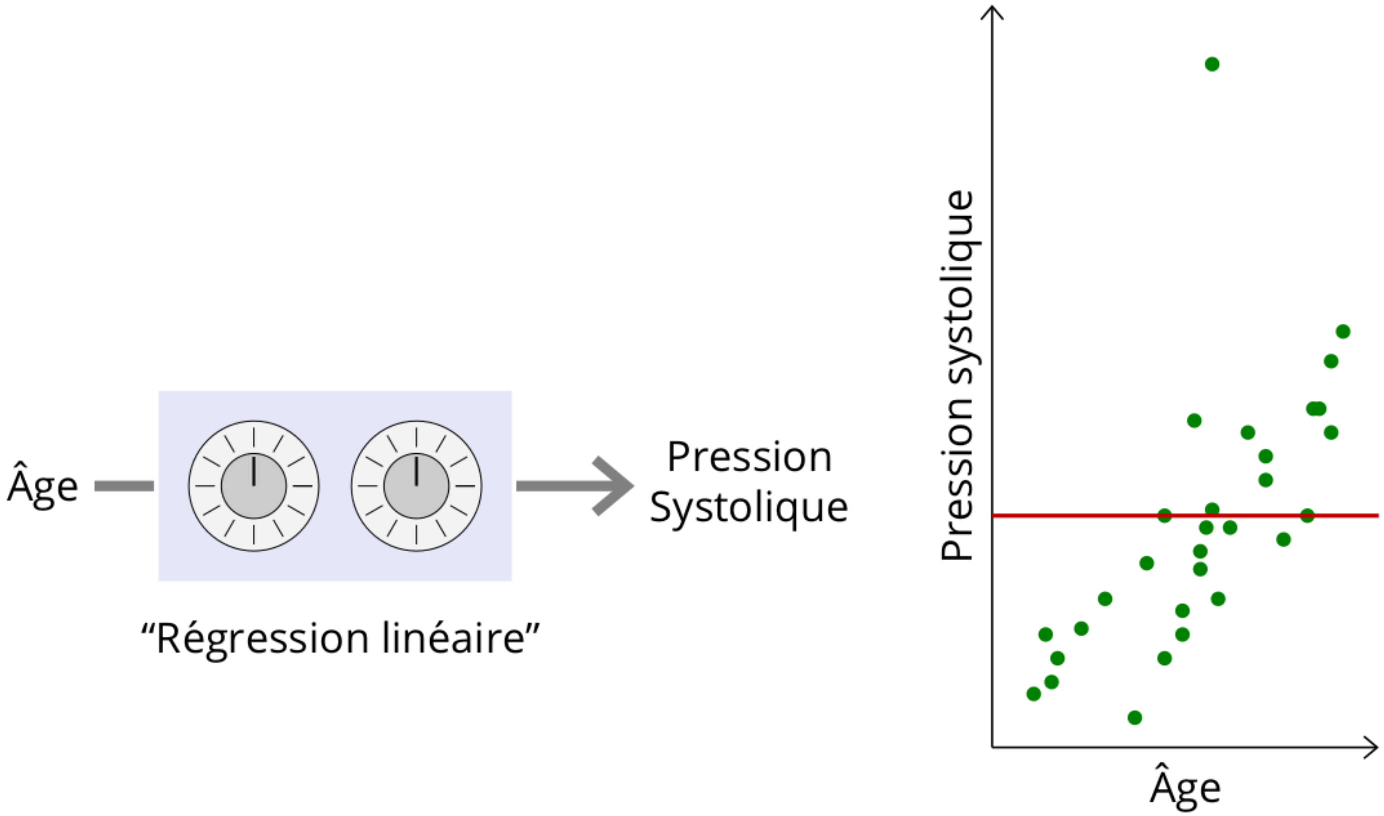
- 1985-1995: The return of neural networks.
- 1988-: Resurgence of probability, focus on uncertainty, general increase in technical depth.
- 1995-2010: New fade of neural networks.
- 2000-: Availability of very large datasets.
- 2010-: Availability of fast commodity hardware (GPUs).
- 2012-: Resurgence of neural networks with deep learning approaches.
- 2017: Attention is all you need (transformers).
- 2022: ChatGPT released to the public.

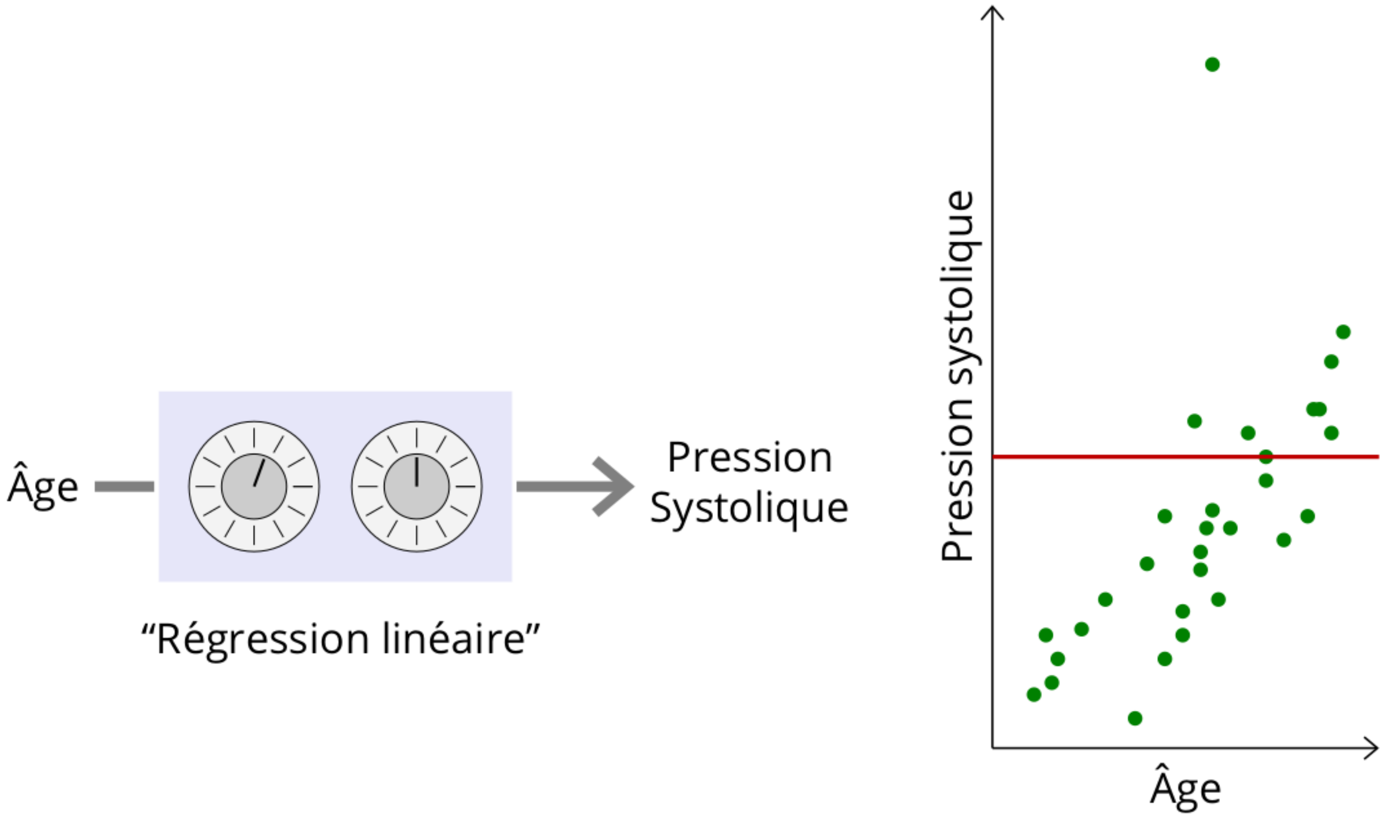
The deep learning revolution



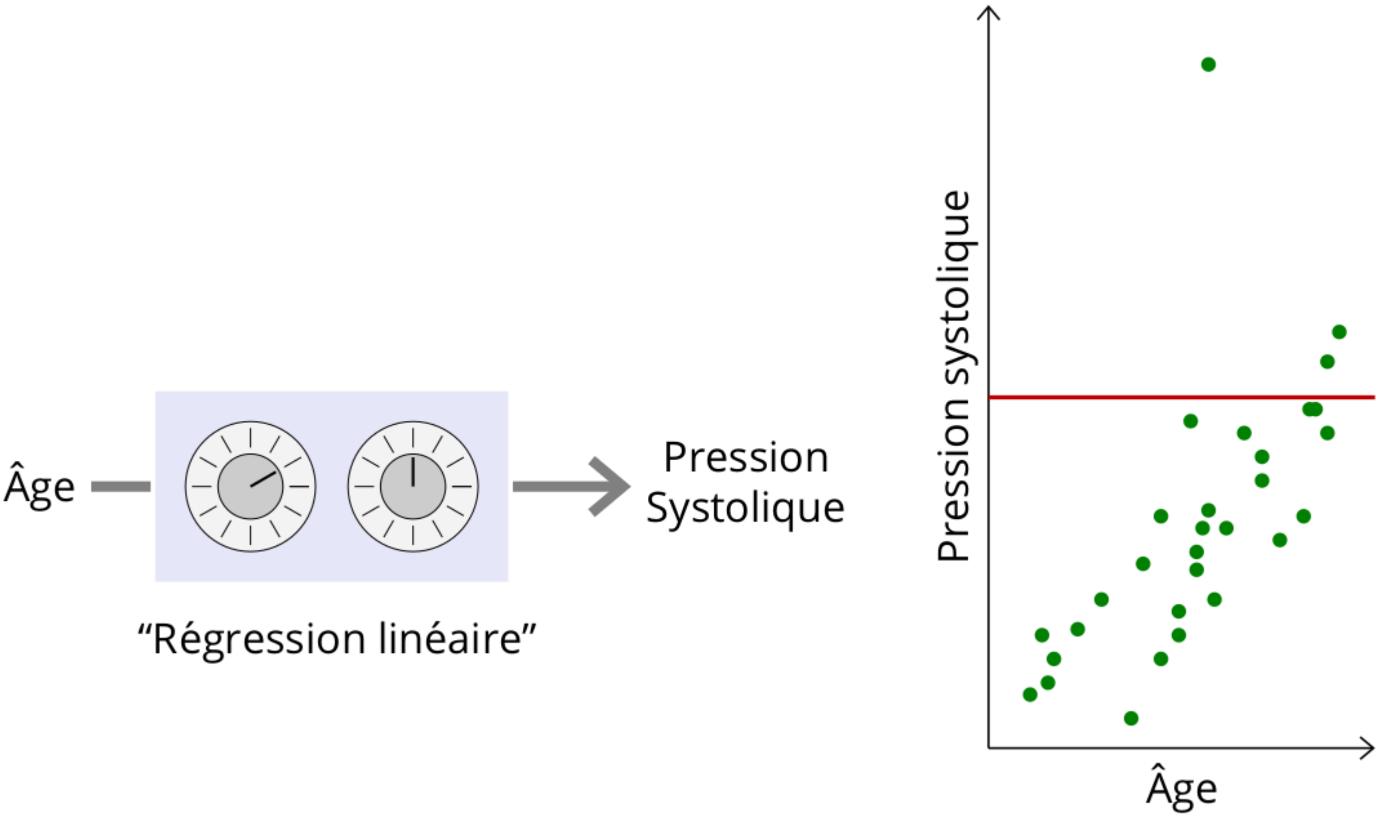
Credits: François Fleuret, 2023.

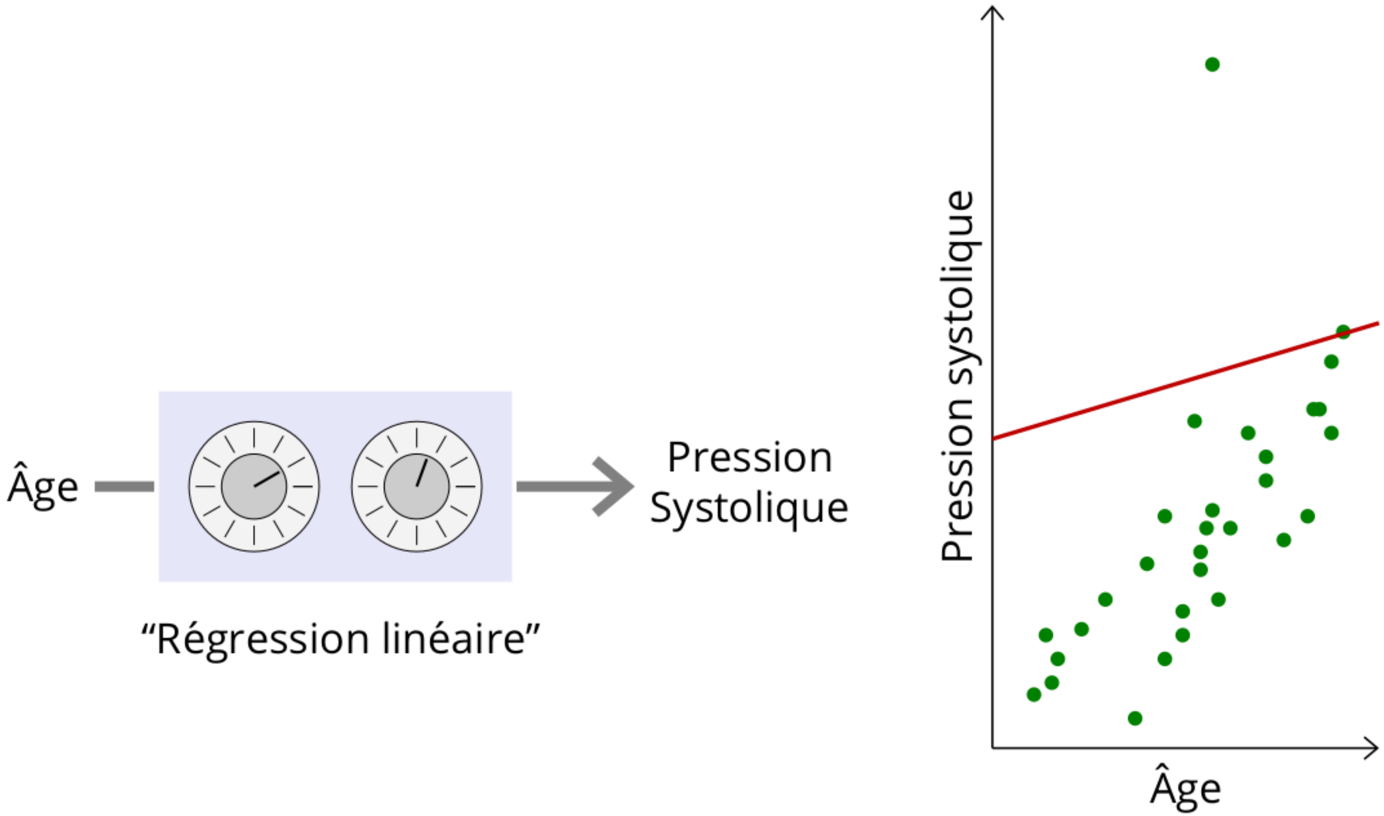

Âge
Programme
Pression
Systolique
Pression systolique
Âge


Âge
“Régression linéaire”
Pression
Systolique
Pression systolique
Âge

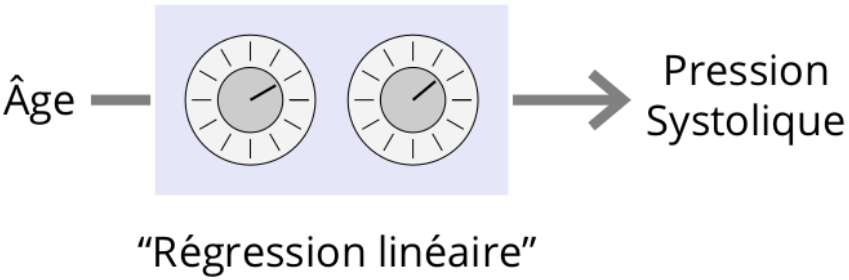


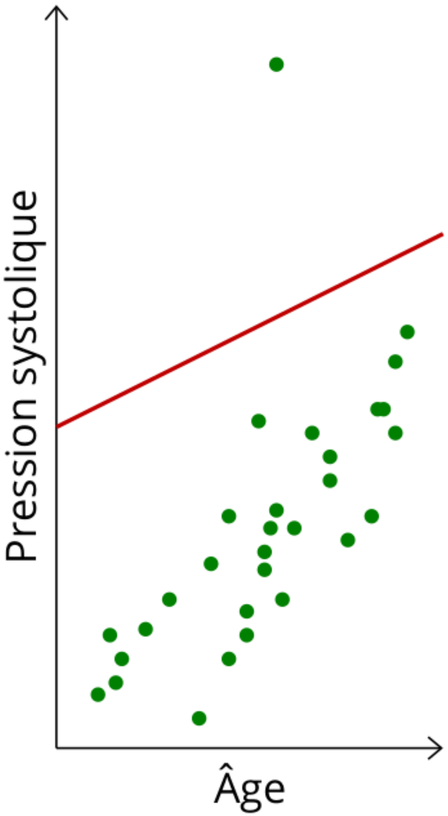
—
Credits: François Fleuret, 2023.

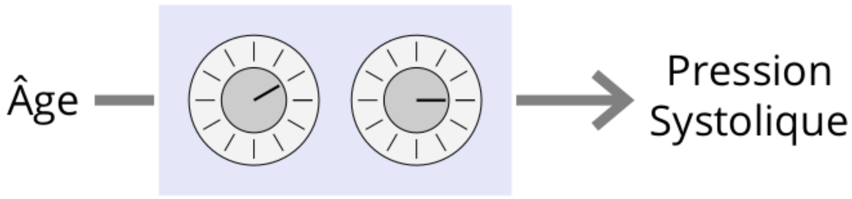

Âge
Pression
Systolique
"Régression linéaire"
Pression systolique
Âge



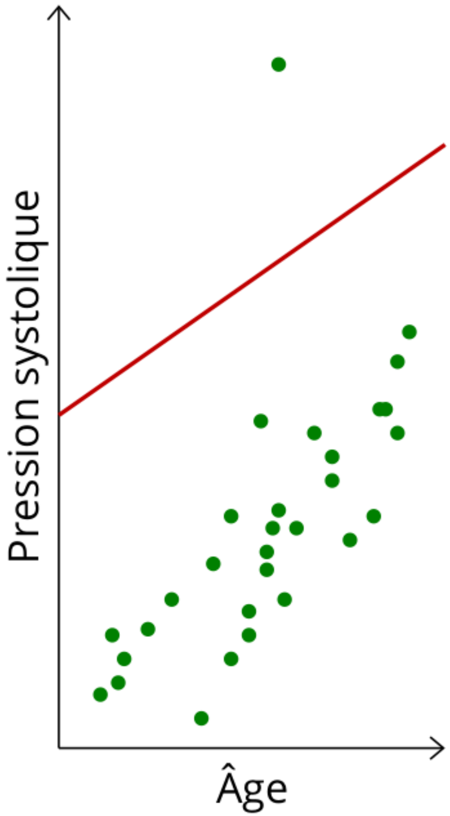
—
Credits: François Fleuret, 2023.

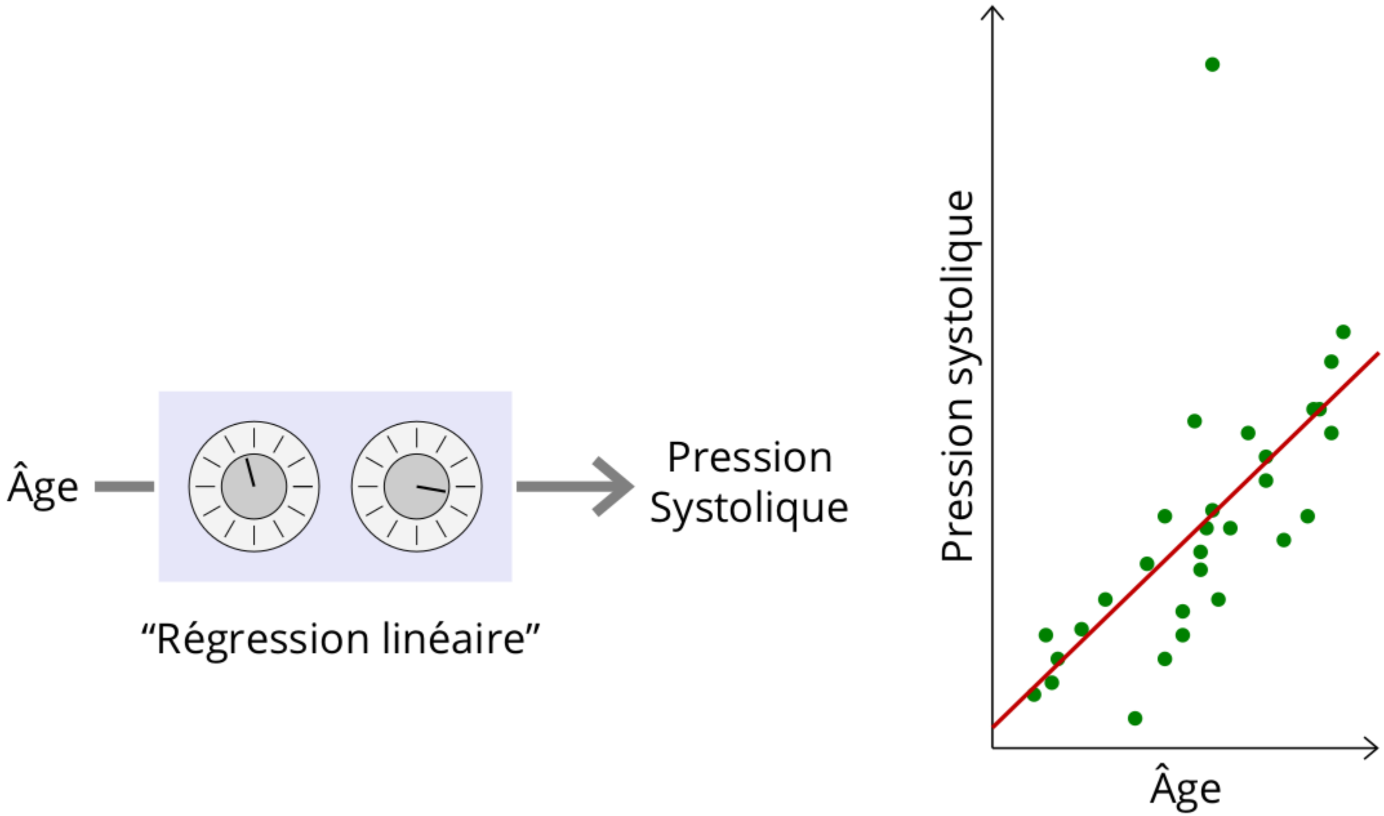

Âge
Pression
Systolique
"Régression linéaire"


Pression systolique
Âge



"Régression linéaire"

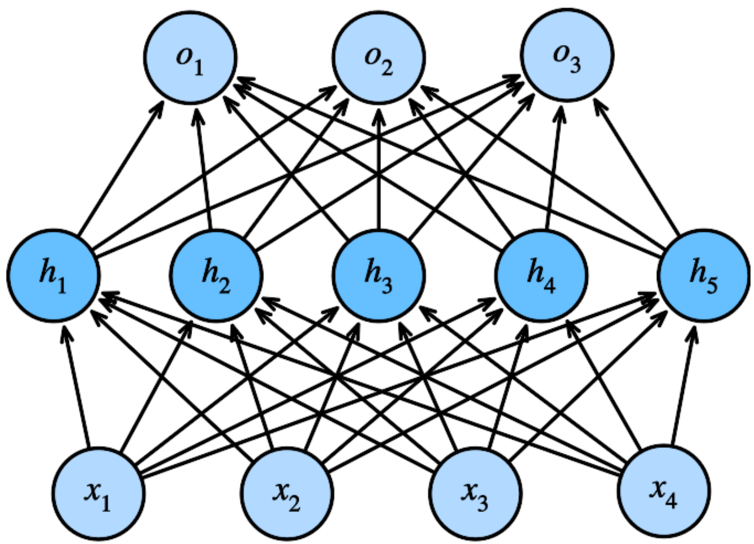


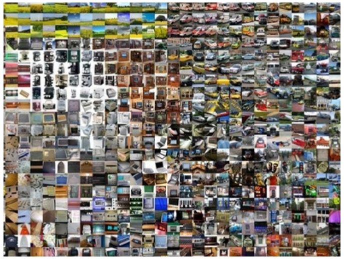


—
Credits: François Fleuret, 2023.

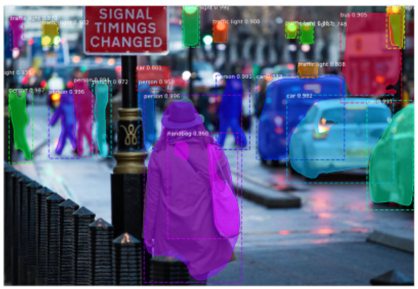
Deep learning **scales up** the statistical and machine learning approaches by

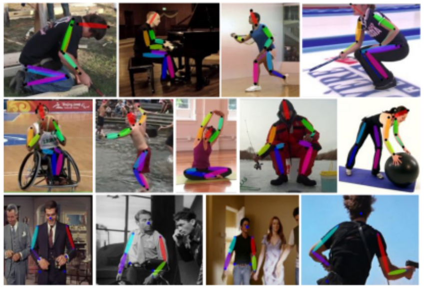
- using larger models known as neural networks,
- training on larger datasets,
- using more compute resources.

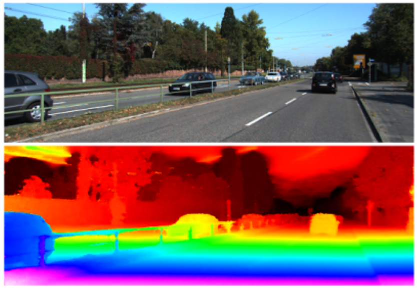


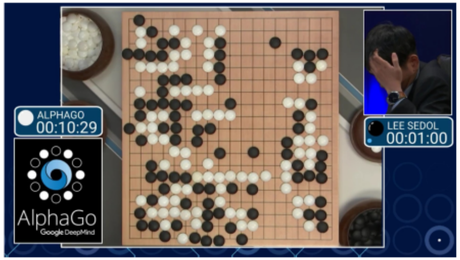


Specialized neural networks can be trained achieve super-human performance on many complex tasks that were previously thought to be out of reach for machines.







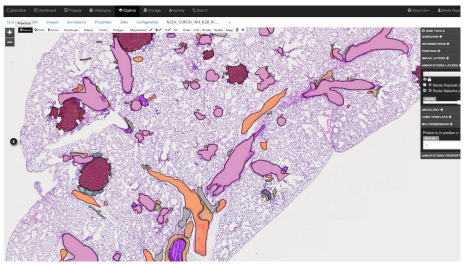


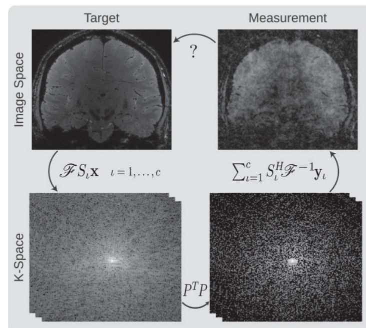
I: Jane went to the hallway.
I: Mary walked to the bathroom.
I: Sandra went to the garden.
I: Daniel went back to the garden.
I: Sandra took the milk there.
Q: Where is the milk?
A: garden

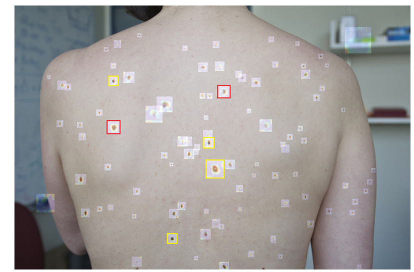
(Top) Scene understanding, pose estimation, geometric reasoning.
(Bottom) Planning, Image captioning, Question answering.

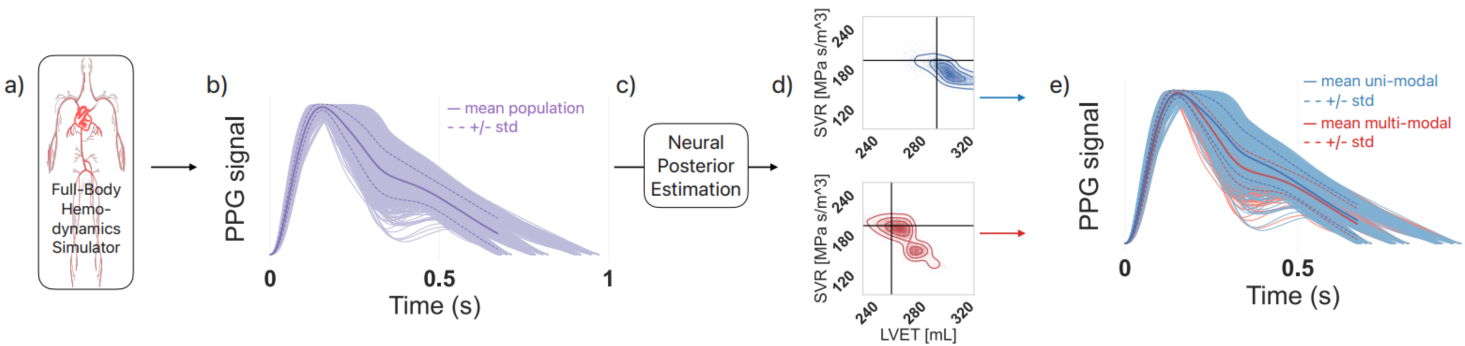
—
Credits: François Fleuret, 2023.

Neural networks form **primitives** that can be transferred to many domains.









(Top) Analysis of histological slides, denoising of MRI images, nevus detection.
(Bottom) Whole-body hemodynamics reconstruction from PPG signals.

The breakthrough

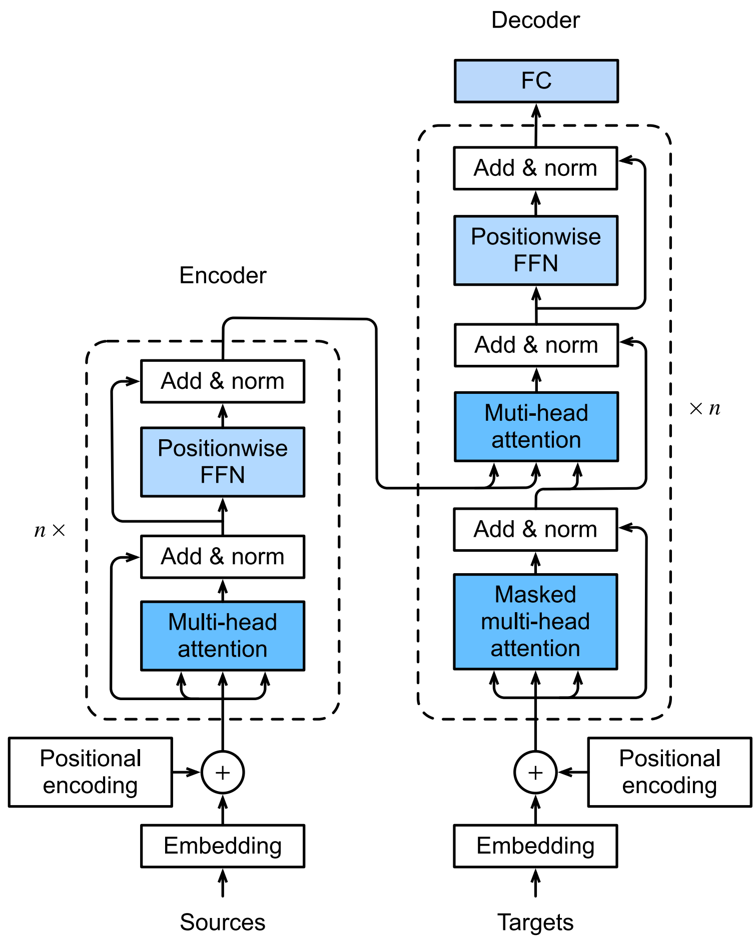
Attention Is All You Need

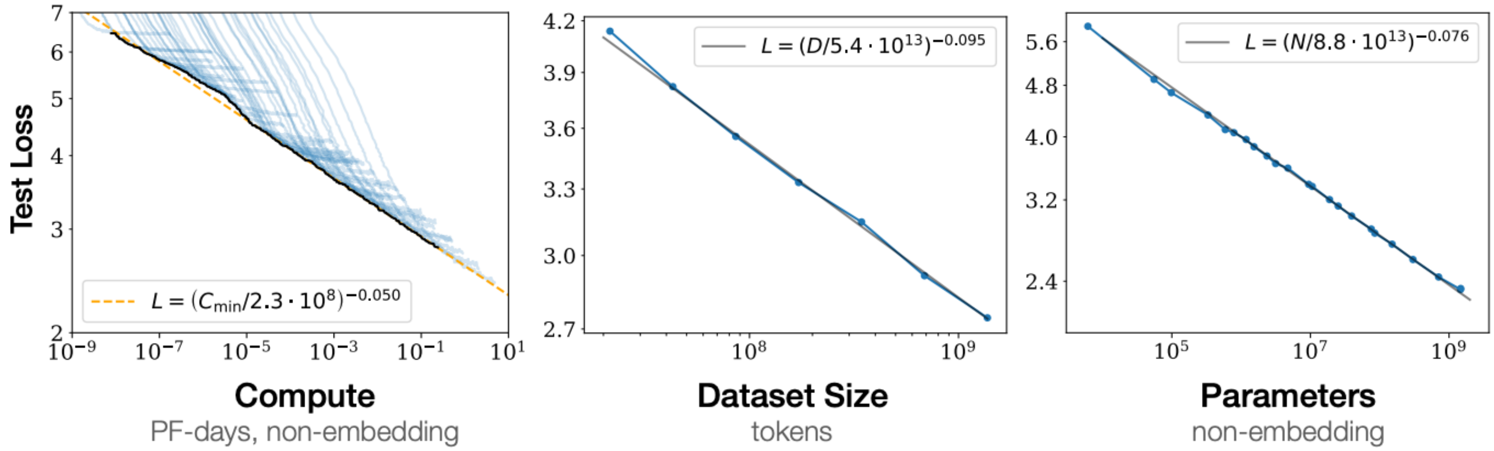
Ashish Vaswani* Google Brain avaswani@google.com
Noam Shazeer* Google Brain noam@google.com
Niki Parmar* Google Research nikip@google.com
Jakob Uszkoreit* Google Research usz@google.com
Llion Jones* Google Research llion@google.com
Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu
Łukasz Kaiser* Google Brain lukaszkaiser@google.com
Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

Vaswani et al., 2017.

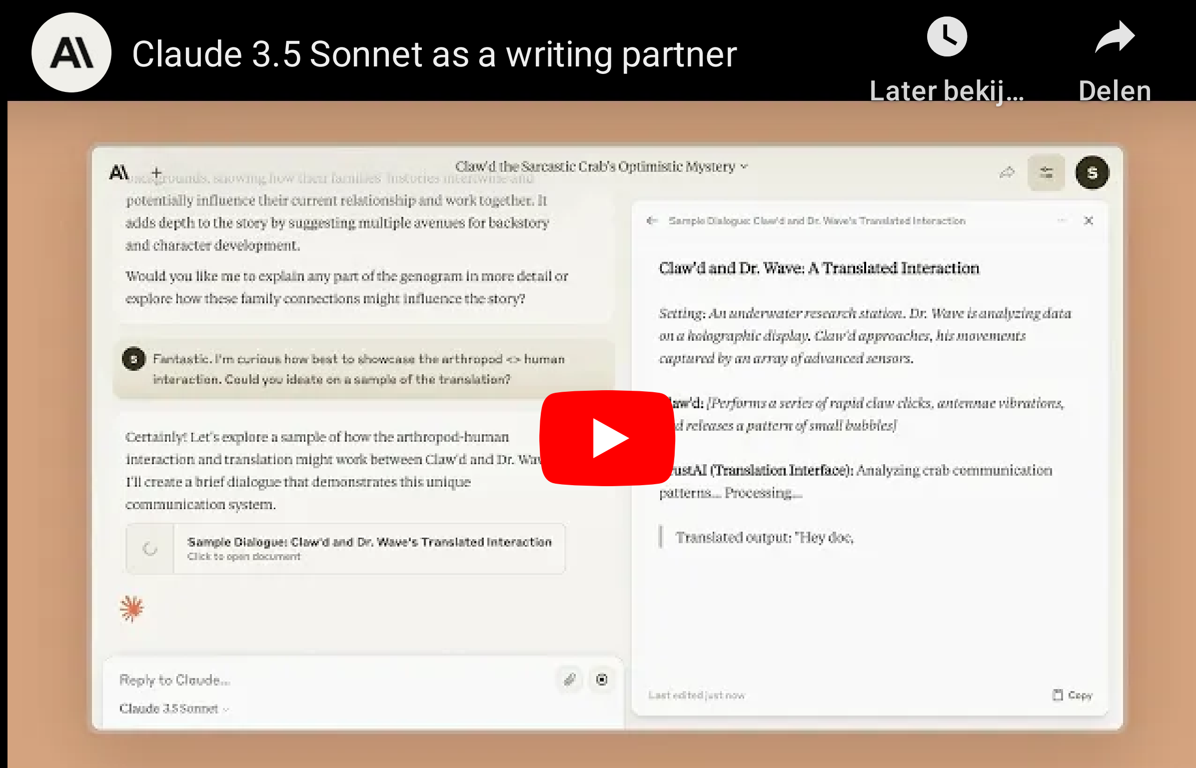




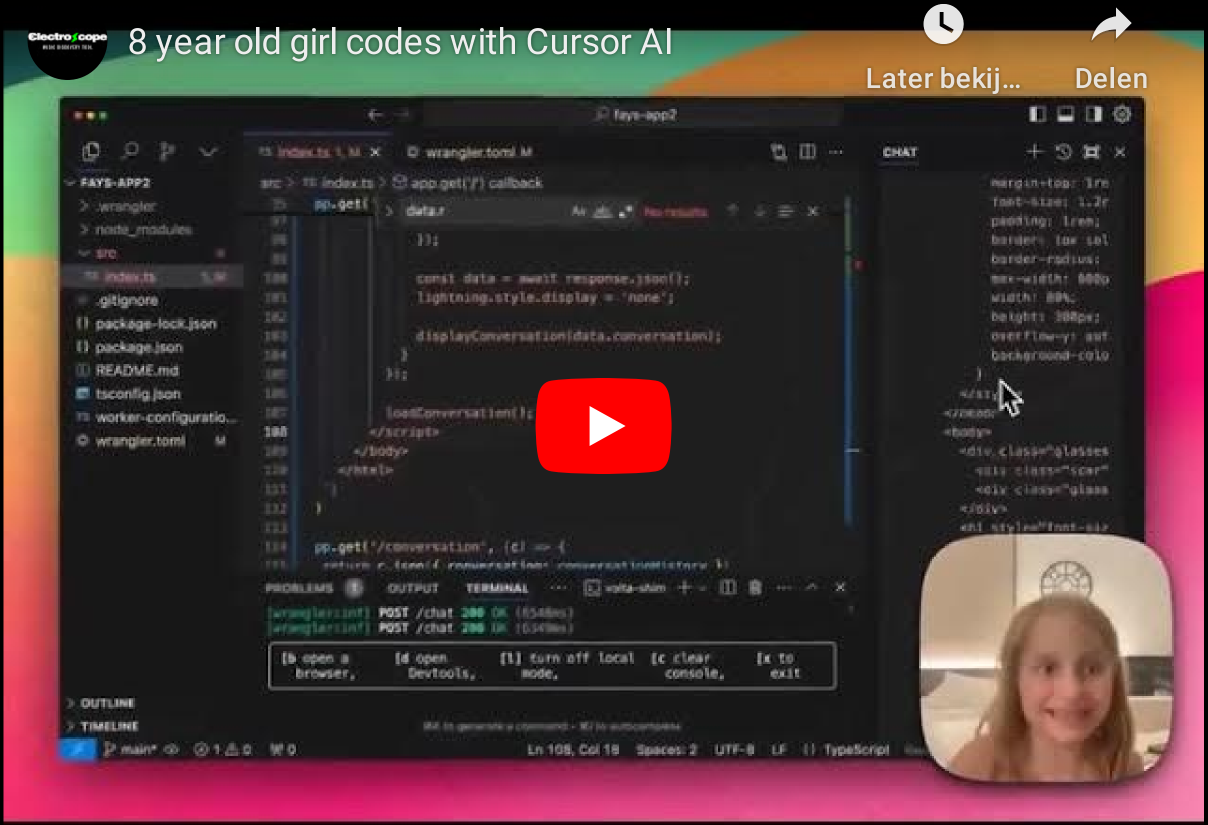
A brutal simplicity:

- The more data, the better the model.
- The more parameters, the better the model.
- The more compute, the better the model.

Scaling up further to gigantic models, datasets, and compute resources keeps pushing the boundaries of what is possible, **with no sign of slowing down**.



Conversational AI assistants (Anthropic, 2024)



Code assistants (Cursor, 2024)

Autonomous cars (Waymo, 2022)

Powering the future of clean energy (NVIDIA, 2023)

How AI is advancing medicine (Google, 2018)

Deep learning can also **solve problems that no one could solve before**.

AlphaFold: From a sequence of amino acids to a 3D structure

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › articles › article

Article | Open access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

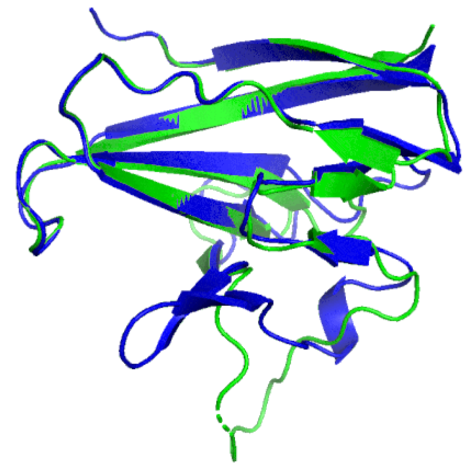
John Jumper ✉, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Jonas Adler, Trevor Back, Stig Petersen, David Reiman, Ellen Clancy, Michal Zielinski, ... Demis Hassabis ✉

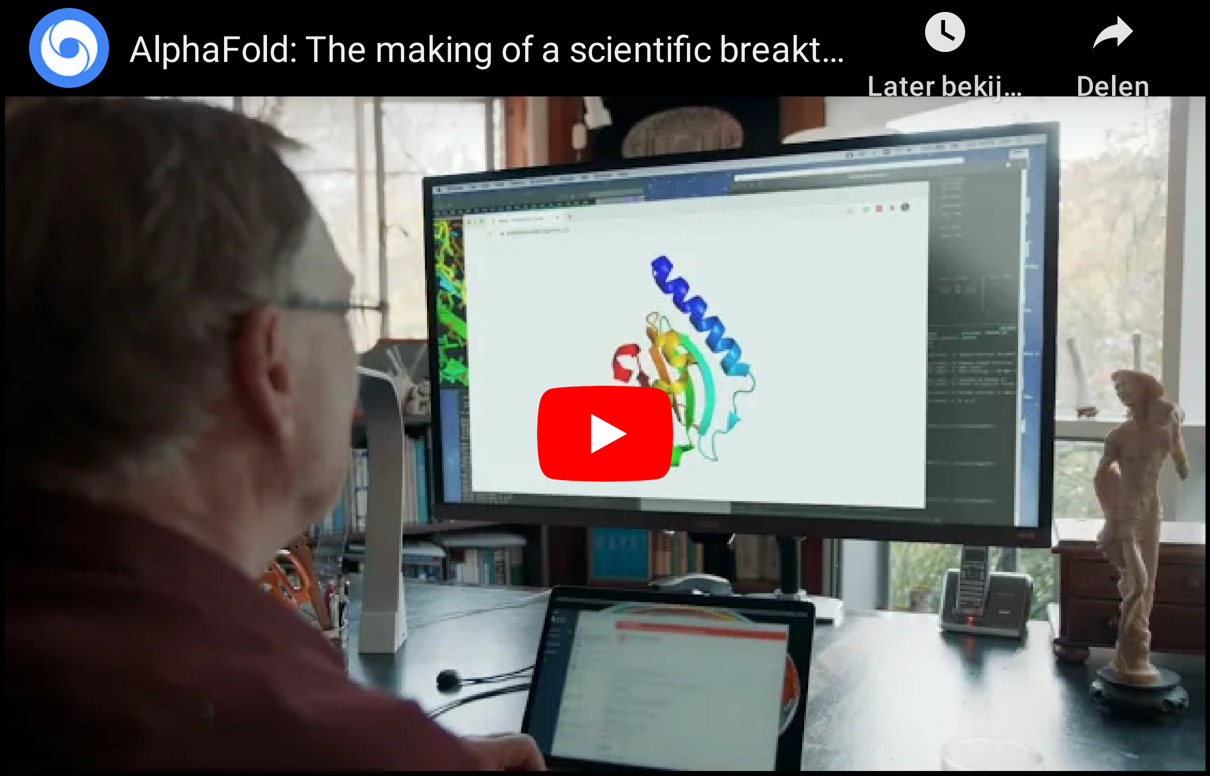
+ Show authors

Nature **596**, 583–589 (2021) | Cite this article

Abstract

Proteins are essential to life, and understanding their structure can facilitate a mechanistic understanding of their function. Through an enormous experimental effort[1,2,3,4], the structures of around 100,000 unique proteins have been determined[5], but this represents a small fraction of the billions of known protein sequences[6,7]. Structural coverage is bottlenecked by the months to years of painstaking effort required to determine a single protein structure. Accurate computational approaches are needed to address this gap and to enable large-scale structural bioinformatics. Predicting the three-dimensional structure that a protein will adopt based solely on its amino acid sequence—the structure prediction component of the 'protein folding problem'[8]—has been an important open research problem for more than 50 years[9]. Despite recent progress[10,11,12,13,14], existing methods fall far short of atomic accuracy, especially when no homologous structure is available. Here we provide the





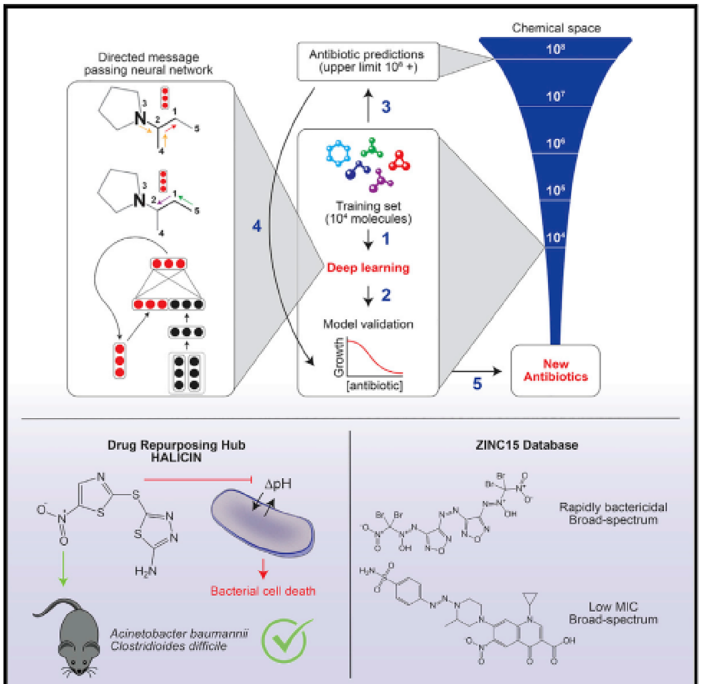
AI for Science (Deepmind, AlphaFold, 2020)

Drug discovery with graph neural networks

Cell

A Deep Learning Approach to Antibiotic Discovery

Graphical Abstract



Authors

Jonathan M. Stokes, Kevin Yang, Kyle Swanson, ..., Tommi S. Jaakkola, Regina Barzilay, James J. Collins

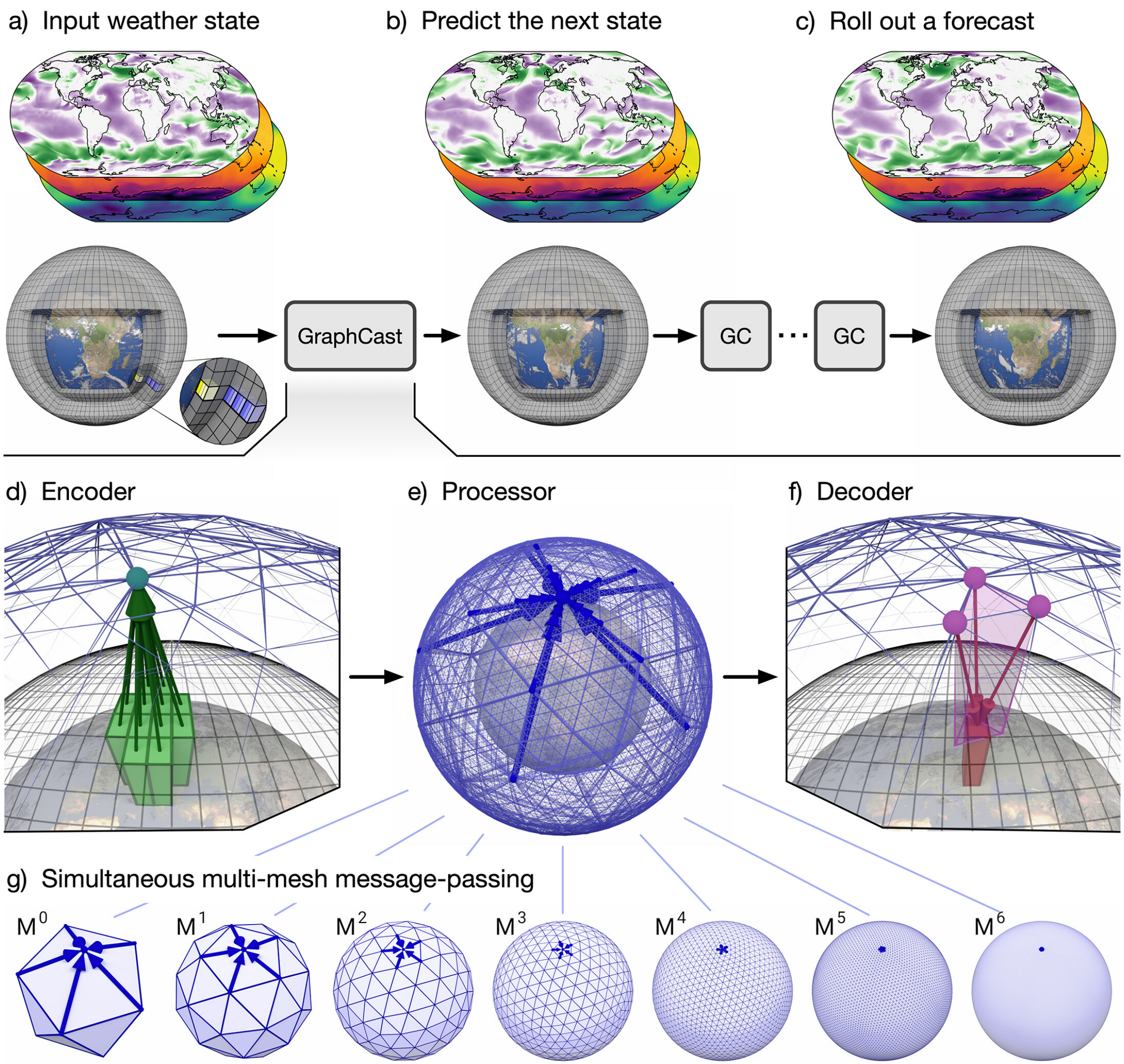
Correspondence

regina@csail.mit.edu (R.B.), jimjc@mit.edu (J.J.C.)

In Brief

A trained deep neural network predicts antibiotic activity in molecules that are structurally different from known antibiotics, among which Halicin exhibits efficacy against broad-spectrum bacterial infections in mice.

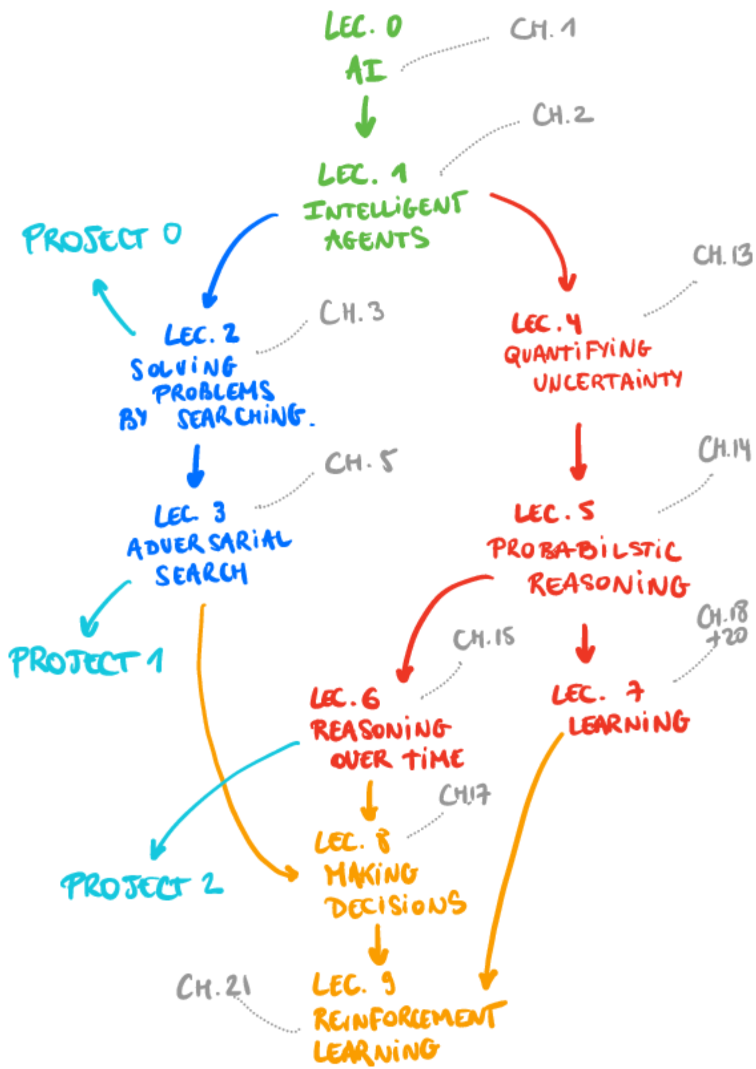
GraphCast: fast and accurate weather forecasts



INFO8006 Introduction to AI

Course outline

- Lecture 0: Artificial intelligence
- Lecture 1: Intelligent agents
- Lecture 2: Solving problems by searching
- Lecture 3: Adversarial search
- Lecture 4: Quantifying uncertainty
- Lecture 5: Probabilistic reasoning
- Lecture 6: Reasoning over time
- Lecture 7: Machine learning and neural networks
- Lecture 8: Making decisions
- Lecture 9: Reinforcement learning


LEC. 0
AI
CH. 1
LEC. 1
INTELLIGENT AGENTS
CH. 2
PROJECT 0
LEC. 2
SOLVING PROBLEMS BY SEARCHING.
CH. 3
LEC. 4
QUANTIFYING UNCERTAINTY
CH. 13
LEC. 3
ADVERSARIAL SEARCH
CH. 5
LEC. 5
PROBABILSTIC REASONING
CH. 14
PROJECT 1
LEC. 6
REASONING OVER TIME
CH. 15
LEC. 7
LEARNING
CH. 18 +20
LEC. 8
MAKING DECISIONS
CH. 17
PROJECT 2
LEC. 9
REINFORCEMENT LEARNING
CH. 21

My mission

By the end of this course, you will have built autonomous agents that efficiently make decisions in fully informed, partially observable and adversarial settings. Your agents will draw inferences in uncertain and unknown environments and optimize actions for arbitrary reward structures.

The models and algorithms you will learn in this course apply to a wide variety of artificial intelligence problems and will serve as the foundation for further study in any application area (from engineering and science, to business and medicine) you choose to pursue.

Goals and philosophy

General

- Understand the landscape of artificial intelligence.
- Be able to write from scratch, debug and run (some) AI algorithms.

Well-established and state-of-the-art algorithms

- Good old-fashioned AI: well-established algorithms for intelligent agents and their mathematical foundations.
- Introduction to materials new from research ($\leq$ 5 years old).
- Understand some of the open questions and challenges in the field.

Practical

- Fun and challenging course projects.

Going further

This course is designed as an introduction to the many other courses available at ULiège and (broadly) related to AI, including:

- INFO8006: Introduction to Artificial Intelligence ← you are there
- DATS0001: Foundations of Data Science
- ELEN0062: Introduction to Machine Learning
- INFO8010: Deep Learning
- INFO8004: Advanced Machine Learning
- INFO9023: Machine Learning Systems Design
- INFO8003: Optimal decision making for complex problems
- INFO0948: Introduction to Intelligent Robotics
- INFO9014: Knowledge representation and reasoning
- ELEN0016: Computer vision

The end.