

Introduction to Deep Learning

Guillaume Baudart

2026

Images generation



<https://platform.openai.com/docs/guides/video-generation>

Images segmentation



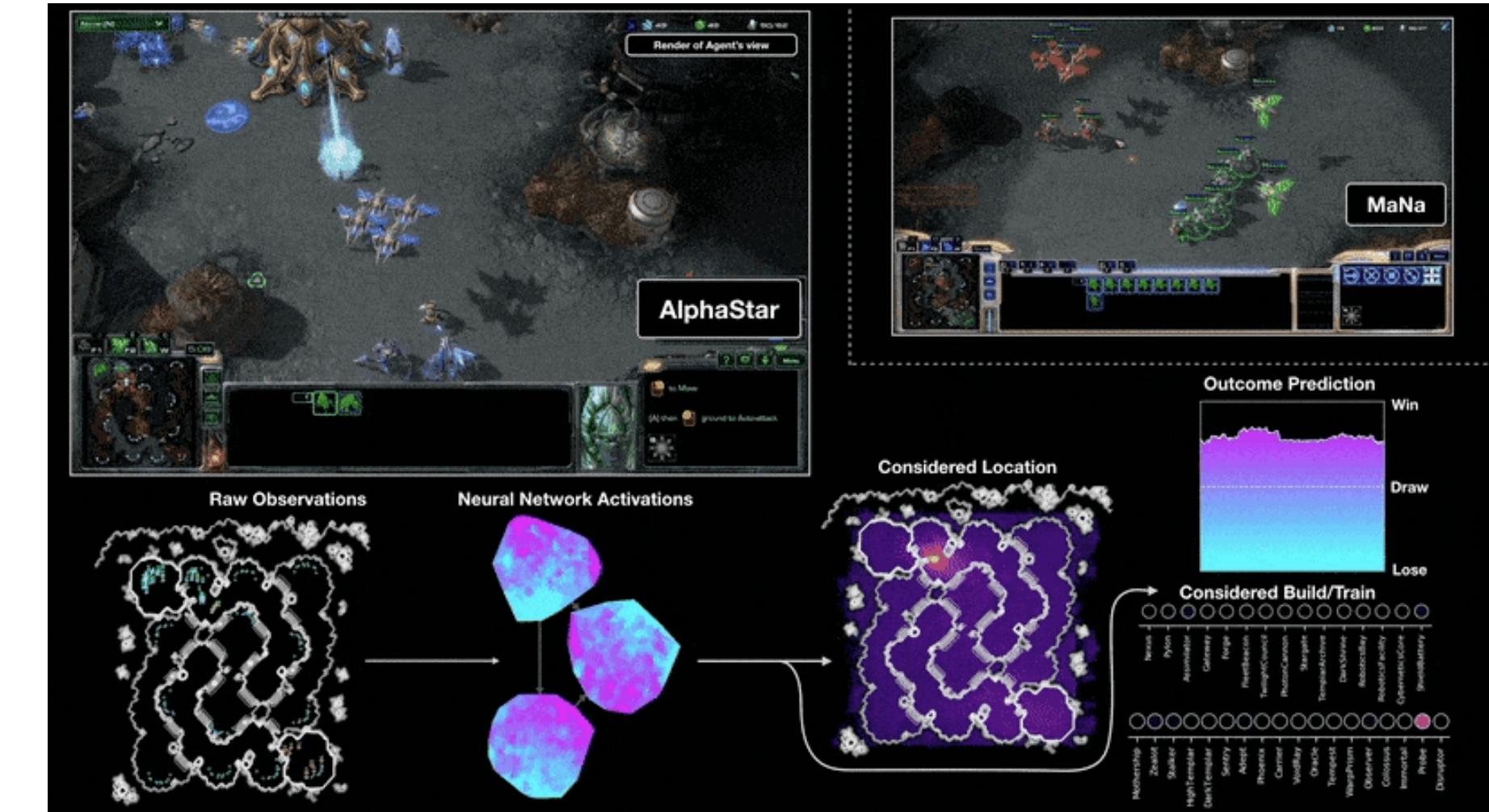
<https://about.fb.com/news/2025/11/new-sam-models-detect-objects-create-3d-reconstructions/>

Autonomous driving



<https://developer.nvidia.com/blog/end-to-end-driving-at-scale-with-hydra-mdp/>

Mastering games



<https://deepmind.google/blog/alphastar-mastering-the-real-time-strategy-game-starcraft-ii/>

What is deep learning?

Artificial Intelligence

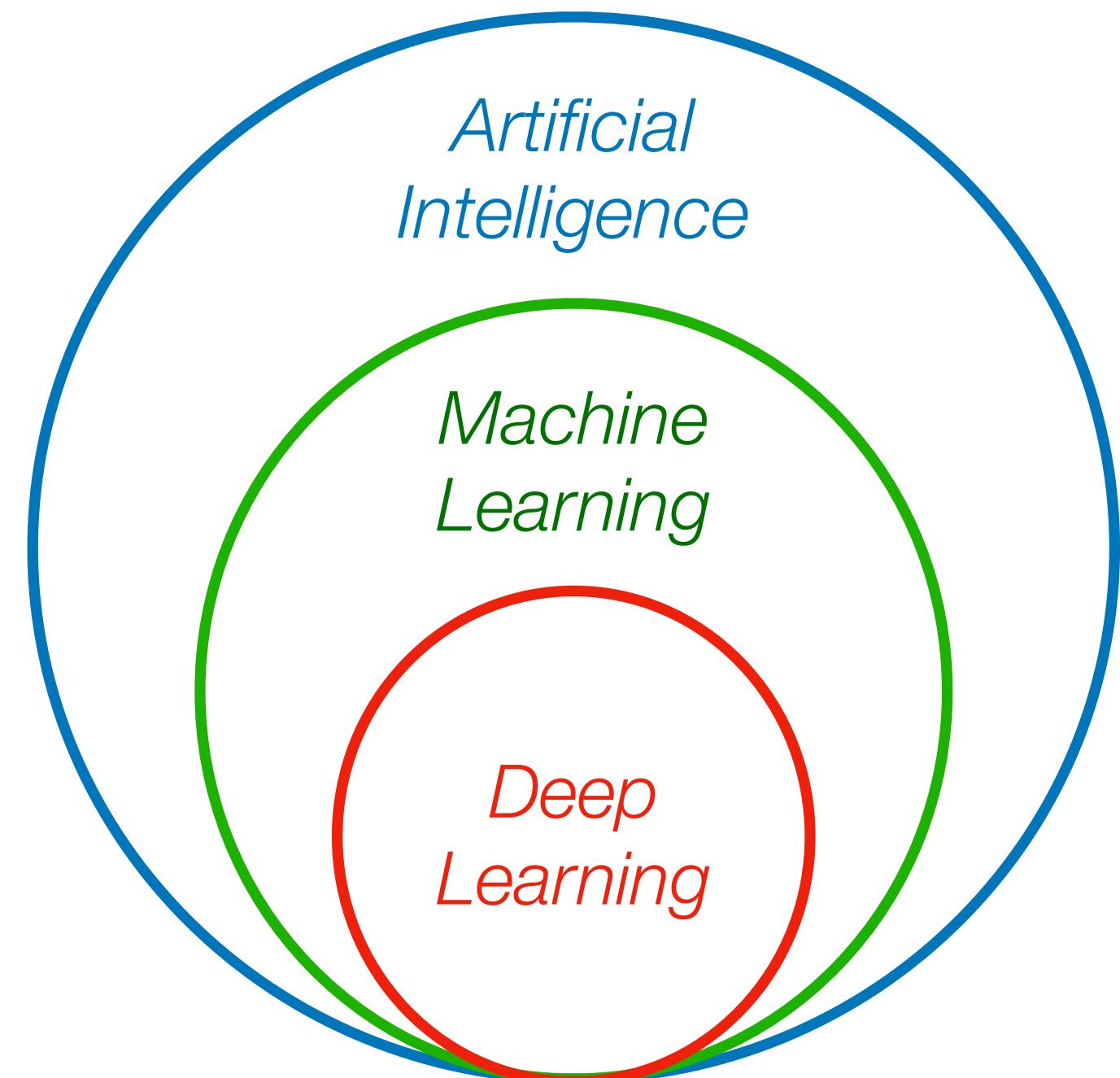
- the capability of computer systems or algorithms to imitate intelligent human behavior

Machine Learning

- a computational method that is a subfield of artificial intelligence and that enables a computer to learn to perform tasks by analyzing a large dataset without being explicitly programmed

Deep Learning

- a form of machine learning in which the computer network rapidly teaches itself to understand a concept without human intervention by performing a large number of iterative calculations on an extremely large dataset



What is deep learning?

Task: image classification Dog vs. Cat



feature extraction

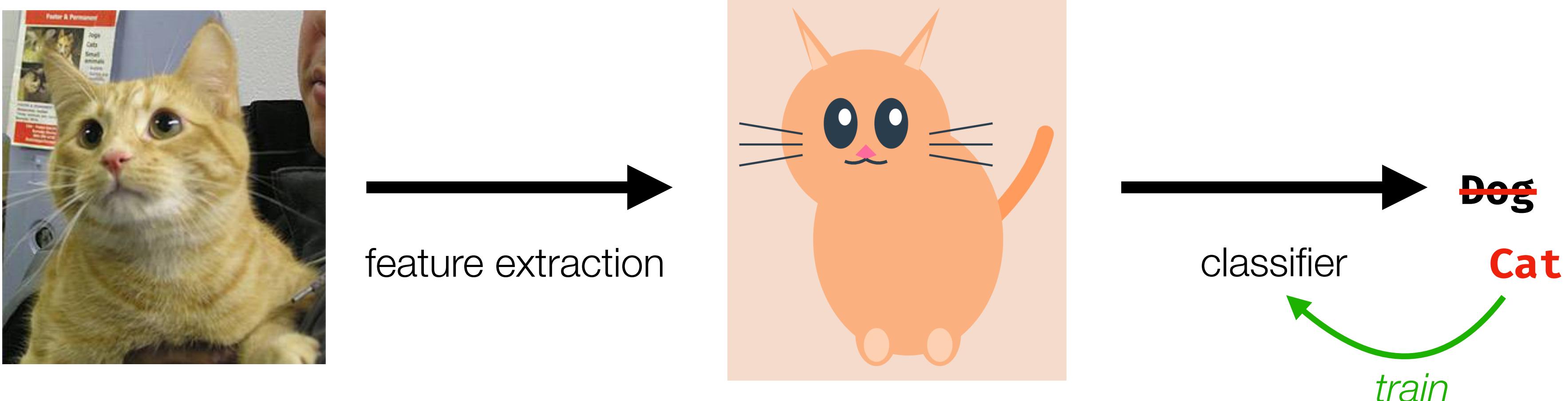


classifier

Dog

What is deep learning?

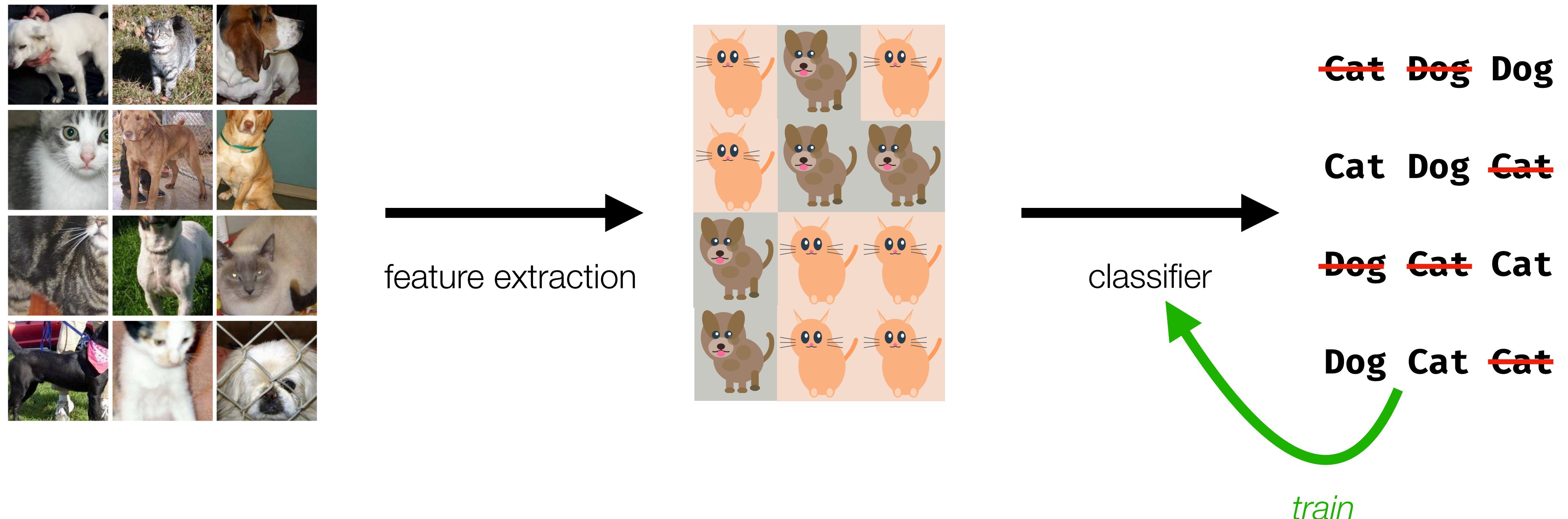
Task: image classification Dog vs. Cat



Supervised machine learning

What is deep learning?

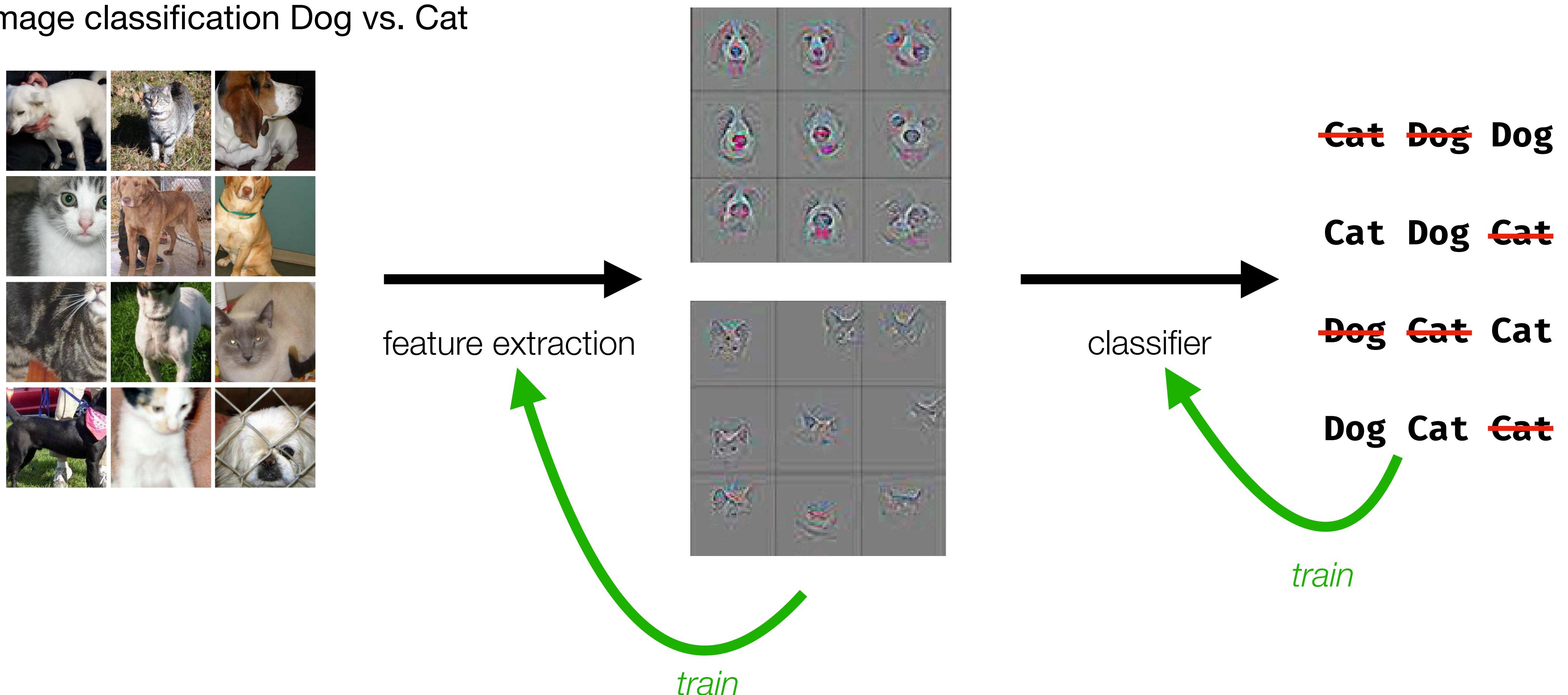
Task: image classification Dog vs. Cat



Supervised machine learning on a complete dataset

What is deep learning?

Task: image classification Dog vs. Cat

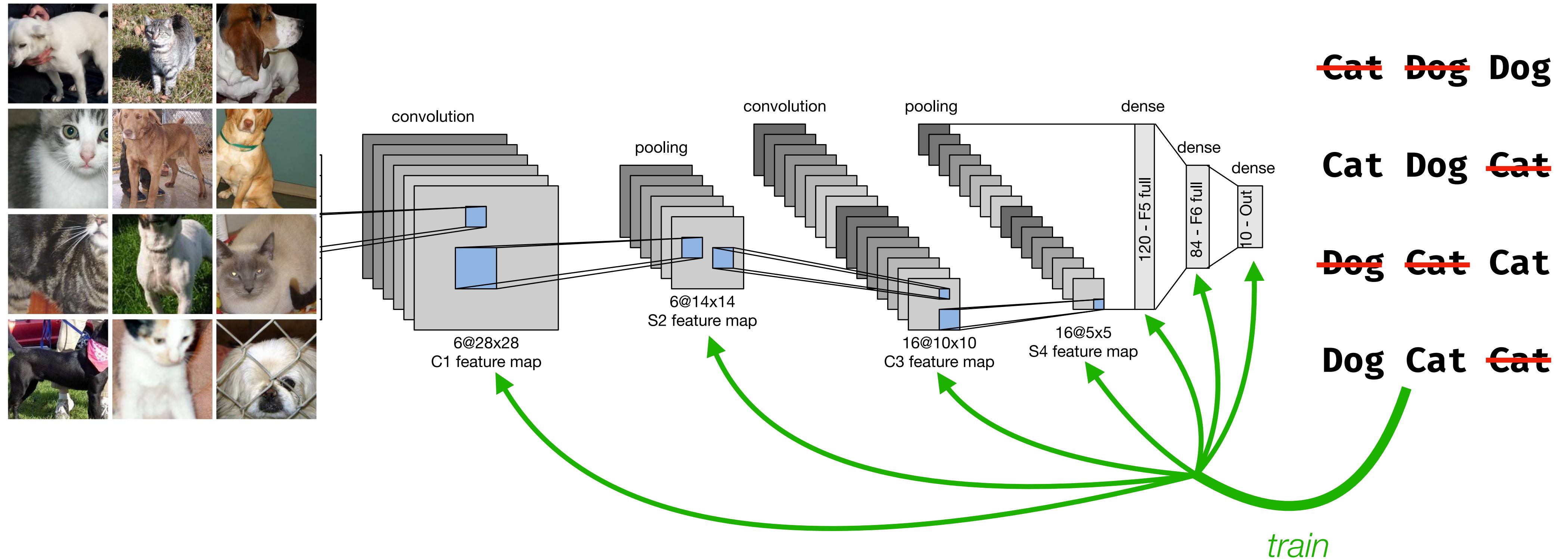


Supervised machine learning on a complete dataset

Deep Learning: learn feature extraction AND classifier with a deep neural network

What is deep learning?

Task: image classification Dog vs. Cat



Supervised machine learning on a complete dataset

Deep Learning: learn feature extraction AND classifier with a deep neural network

Many flavors of deep learning

What are the data?

- Image
- Video
- Sound
- Tabular
- Text
- ...

What is the task?

- Classification
- Recommandation
- Generation
- Denoising
- Translation
- Segmentation
- ...

Supervised learning

- Data: (x_i, y_i)
- Task: predict y for a new x

Unsupervised learning

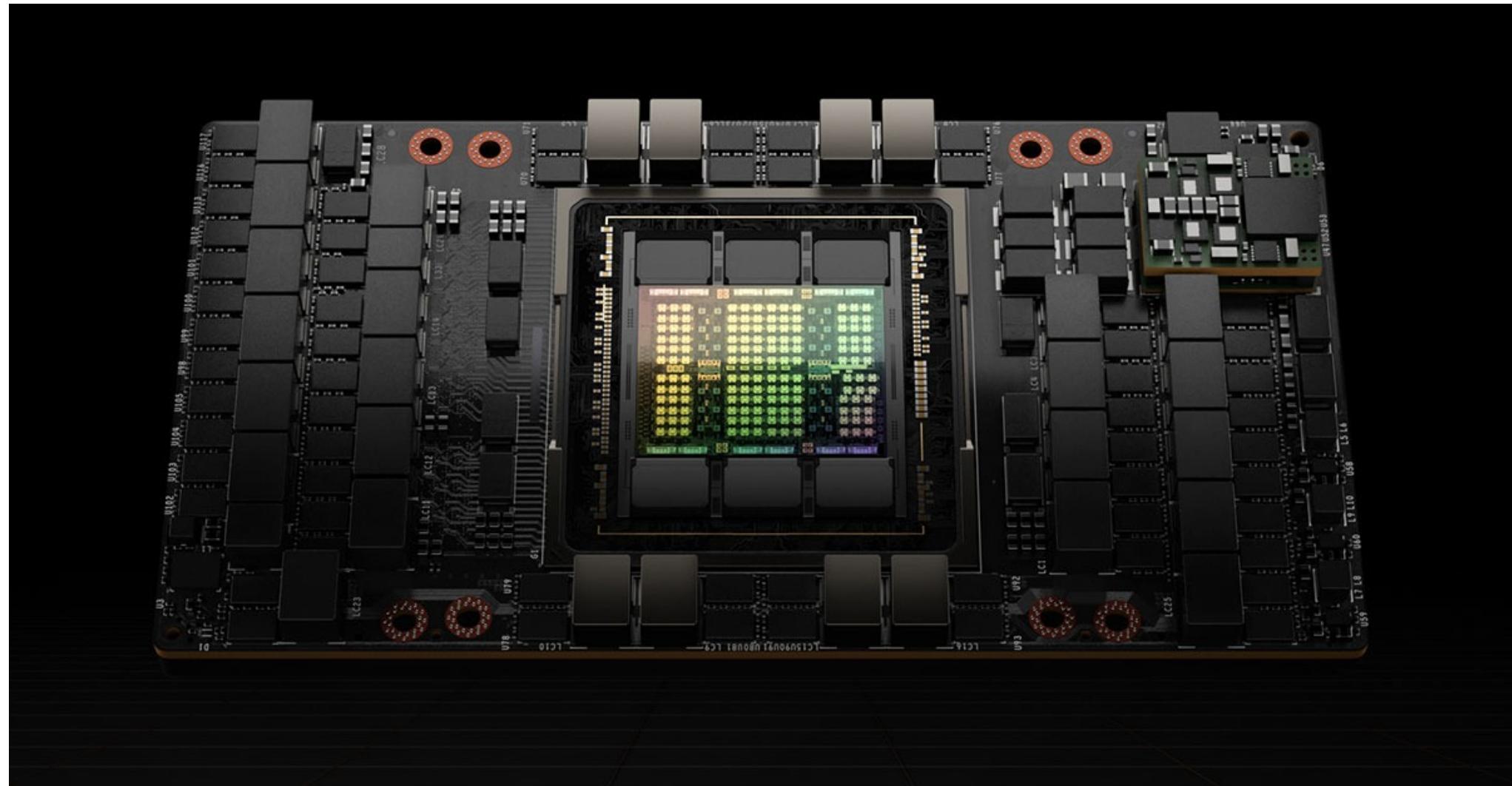
- Data: (x_i)
- Task: generate new x similar to (x_i)

Reinforcement learning

- Data: interaction with an environment (s_i, a_i, r_i)
- Sequence of state, action, reward
- Task: learn policy to maximize cumulative reward

What makes it possible?

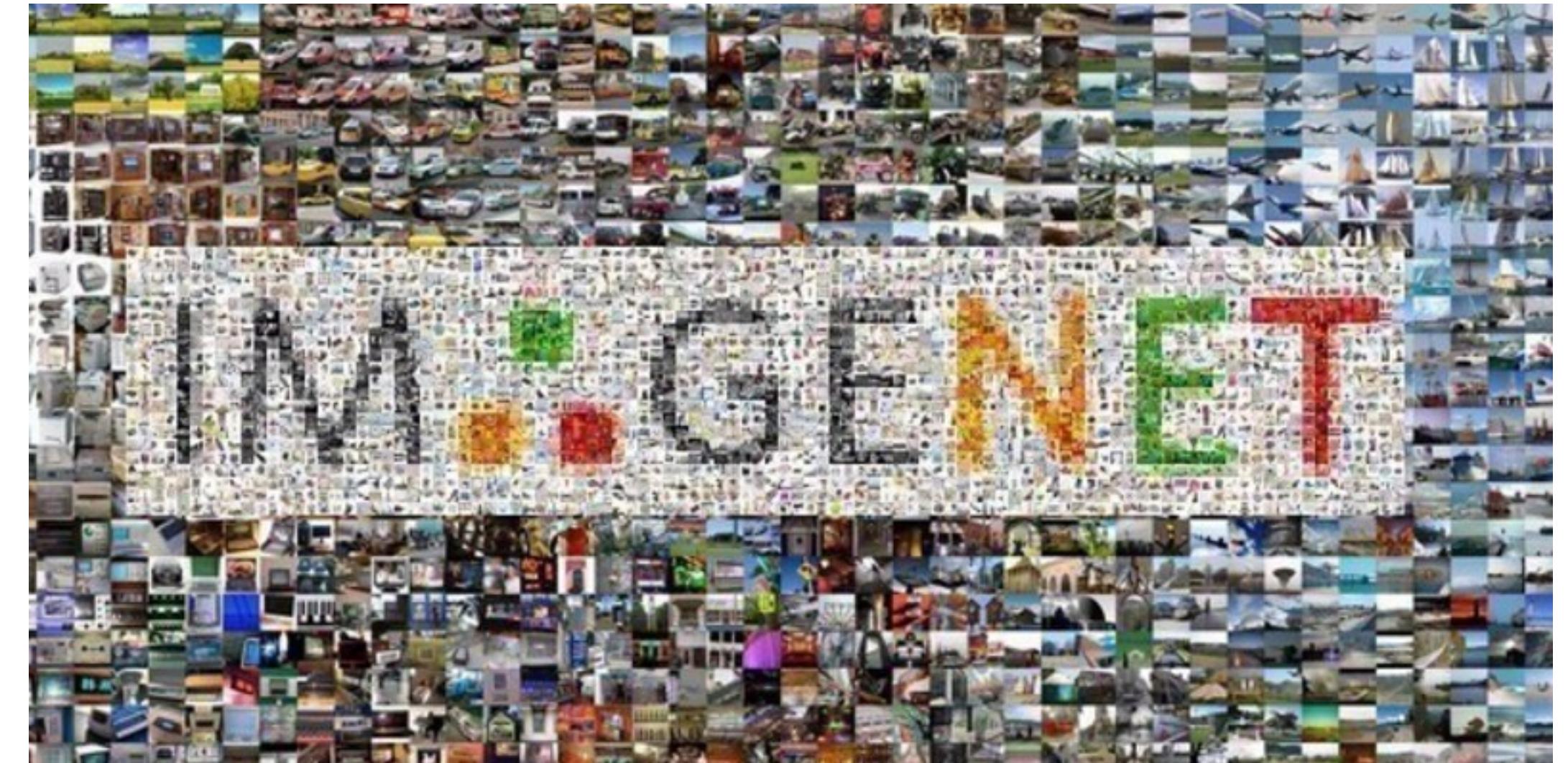
Massive compute



NVIDIA H100 GPU

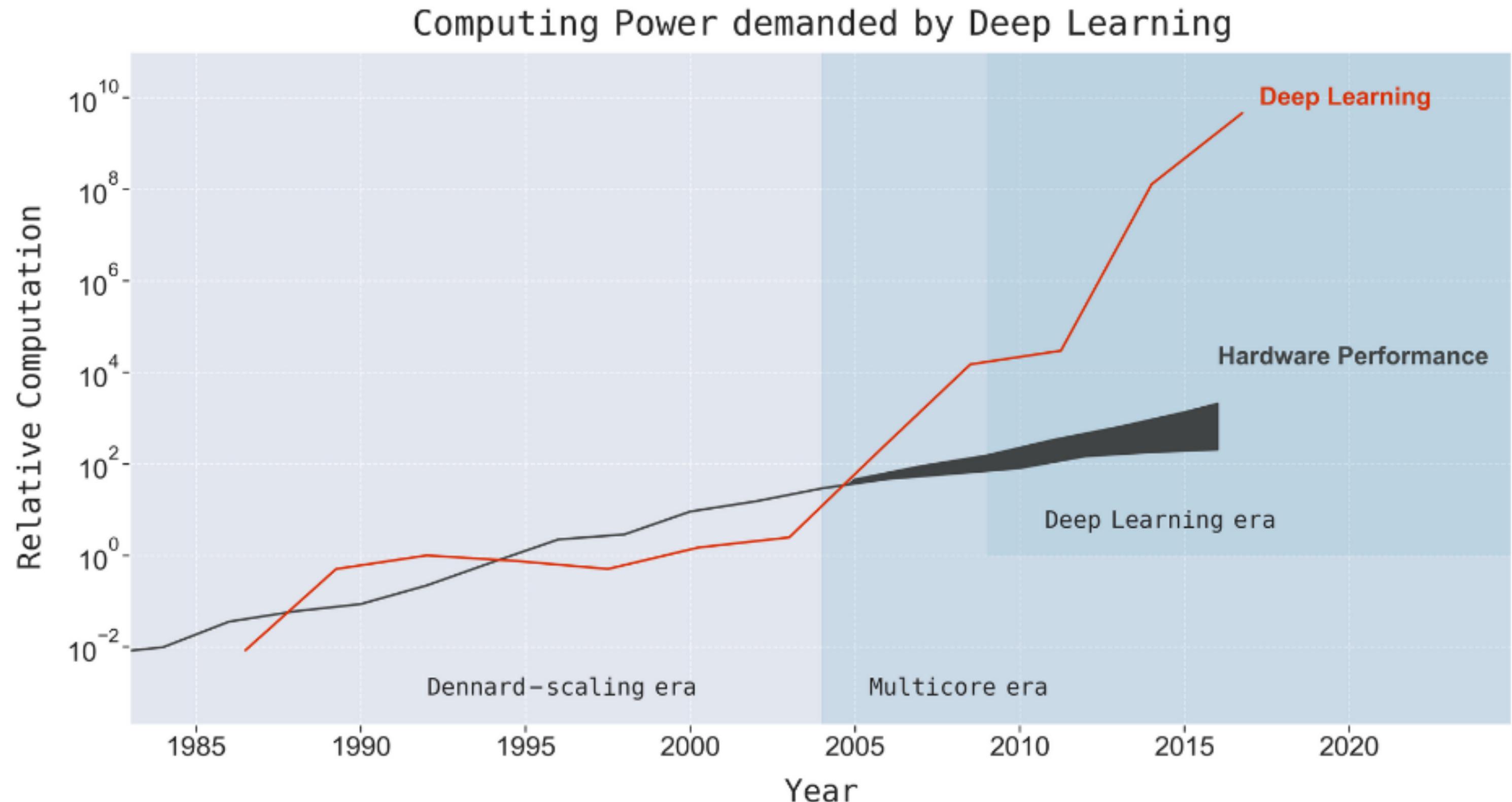
xAI Colossus (2024): 150,000 H100

Very (very) large datasets



14 million hand annotated images, 20 000 classes

Deep learning evolution



A few examples:

- 2012, AlexNet 2 GPUs 5-6 days
- 2017, ResNeXt-101, 8 GPUs 10 days
- 2019, NoisyStudent 1000 TPUs for 6 days

Deep Learning evolution

Training compute (FLOPs) of milestone Machine Learning systems over time

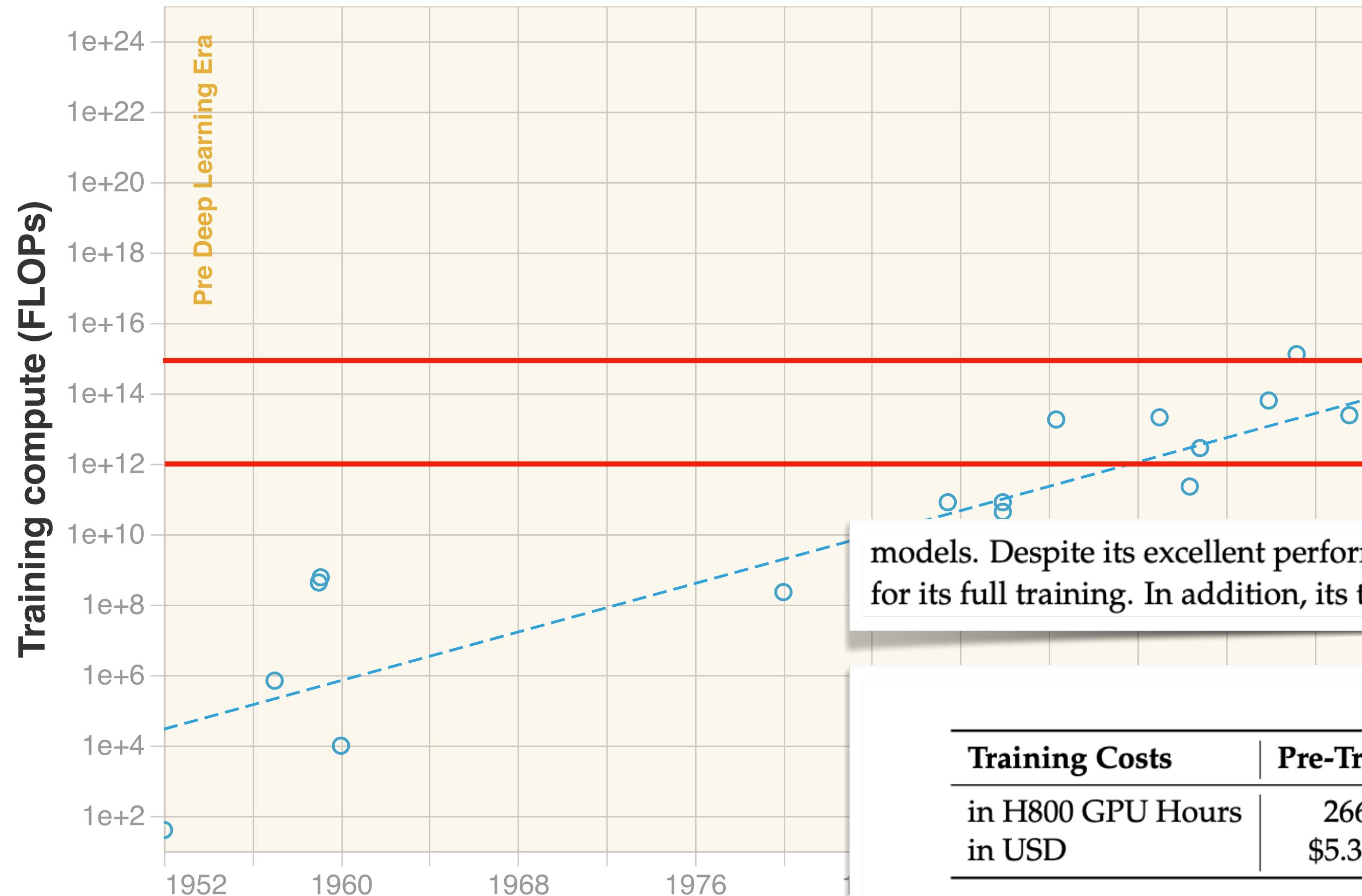
n = 121



Deep Learning evolution

Training compute (FLOPs) of milestone Machine Learning systems over time

n = 121



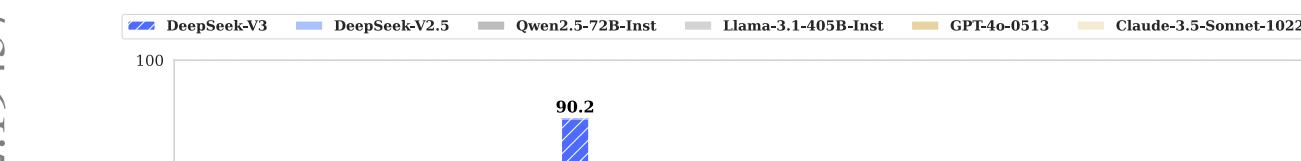
DeepSeek-V3 Technical Report

DeepSeek-AI

research@deepseek.com

Abstract

We present DeepSeek-V3, a strong Mixture-of-Experts (MoE) language model with 671B total parameters with 37B activated for each token. To achieve efficient inference and cost-effective training, DeepSeek-V3 adopts Multi-head Latent Attention (MLA) and DeepSeekMoE architectures, which were thoroughly validated in DeepSeek-V2. Furthermore, DeepSeek-V3 pioneers an auxiliary-loss-free strategy for load balancing and sets a multi-token prediction training objective for stronger performance. We pre-train DeepSeek-V3 on 14.8 trillion diverse and high-quality tokens, followed by Supervised Fine-Tuning and Reinforcement Learning stages to fully harness its capabilities. Comprehensive evaluations reveal that DeepSeek-V3 outperforms other open-source models and achieves performance comparable to leading closed-source models. Despite its excellent performance, DeepSeek-V3 requires only 2.788M H800 GPU hours for its full training. In addition, its training process is remarkably stable. Throughout the entire training process, we did not experience any irrecoverable loss spikes or perform any rollbacks. The model checkpoints are available at <https://github.com/deepseek-ai/DeepSeek-V3>.



Training Costs	Pre-Training	Context Extension	Post-Training	Total
in H800 GPU Hours	2664K	119K	5K	2788K
in USD	\$5.328M	\$0.238M	\$0.01M	\$5.576M

Table 1 | Training costs of DeepSeek-V3, assuming the rental price of H800 is \$2 per GPU hour.

Deep Learning Frameworks

Support for vectorized computations, automatic differentiation, and hardware accelerators

- 1988 Mathematica (Wolfram)
- 2002 Torch (Idiap)
- 2007 Theano (Université de Montréal)
- 2015 Tensorflow (Google Brain)
- 2016 PyTorch (Facebook)
- 2018 JAX (Google)



Hands-on: Transfer learning with VGG

Introduction to deep learning

Course content

Understand deep learning models

- How they work
- How they are trained
- How they are used

Main topics

- Deep learning pipeline
- Loss and optimization
- Training and running on GPUs
- Convolutional neural networks (CNN)
- Recurrent neural networks (RNN)
- Language models, Attention, and transformers
- Agentic programming

Organization

11 courses

- 2h session: Friday 14h-16h
- Mix course / practicals

Bring your machines!

Hardware

- Personal computers
- Google colab GPUs
- TBD...

Courses / books

- Lelarge, M. Dataflowr
- Prince, S., Understanding Deep Learning
- Zhang, A. et al., Dive into Deep Learning

Evaluation

Project: 70% + Exam: 30%

Project: experiment with existing projects

- Small groups
- Understand a paper/technique/code
- Add experiments
- 30mn presentation (including Q&A)

Exam: as usual...

Using LLMs

Rules:

- All text and code you submit is your responsibility
- No plagiarism
- No external tools during the exam

When using LLM exercise caution and critical thinking

- Is this answer true? Correct? Optimal?
- Do I understand the answer? Could I find it myself without LLM help?