Deep Learning

Course Overview

Deep learning

- Basics
  - Overfitting and underfitting
  - Training framework
  - Back propagation
- Modern neural network
  - Convolutional neural network
  - Recurrent neural network/LSTM
  - Transformer
  - GNN?
- Reinforcement learning
- Unsupervised/self supervised learning

Application

- Computer vision
  - Segmentation
  - Object detection
- Generative AI
  - Autoencoder
  - Generative Adversarial Network
  - LLM, multimodal LLM
  - Diffusion models
- Implementation and real time execution issues
  - Model compression and acceleration
    - Pruning, quantization, low complexity models
    - For LLM
- Others
  - Latest trends will be added anytime
  - Application example to EE fields if time admitted

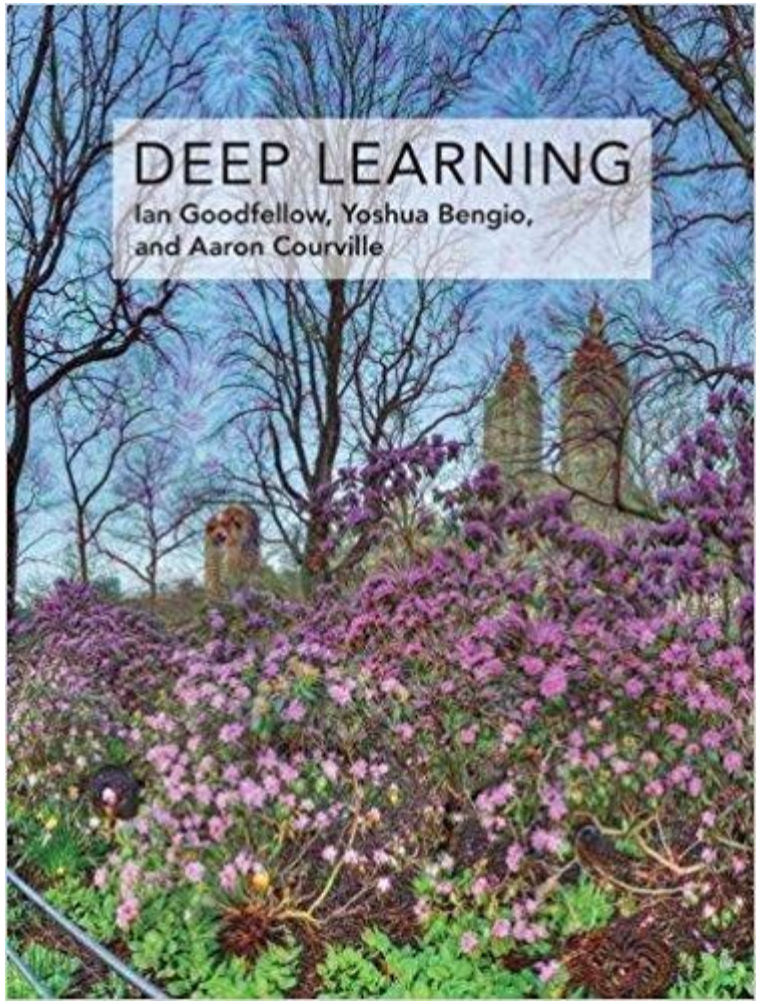
Target Audience

- Background to have
  - Linear algebra
  - Probability and statistics
  - Programming, especially python
  - Nice to know about machine learning

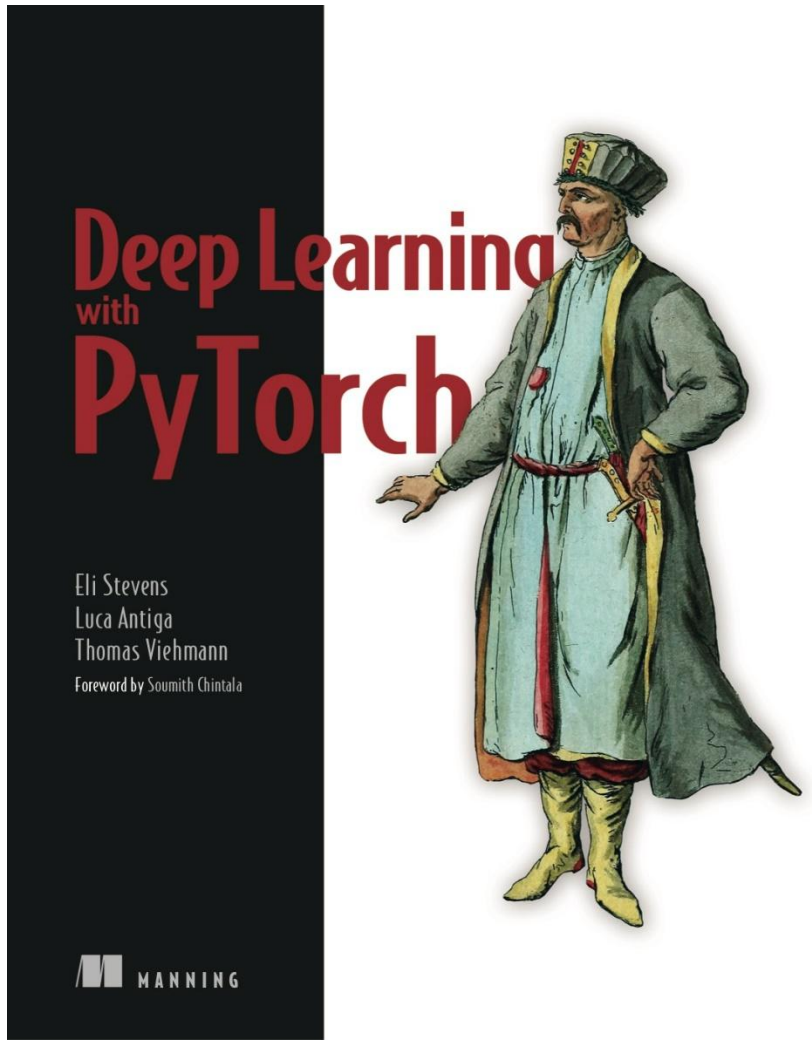
Administrative

- 1.學期作業、考試、評量
  - Lab x5 or x4 + 1 midterm (60%)
    Final project (20%) (1 or 2 persons in a group. topic survey + implementation, TBA)
    Final exam (20%)
- 2.教學方法及教學相關配合事項（如助教、網站或圖書及資料庫等）
  - all materials will be at NYCU e3 site
  - 上課以實體為主
    - 上課錄影 (on youtube, TBA)
  - 因應人數控管，一律線上選課，不手動加簽。電機系電子所優先選課，產創學院如未選到課，可洽所辦登記保留名額。

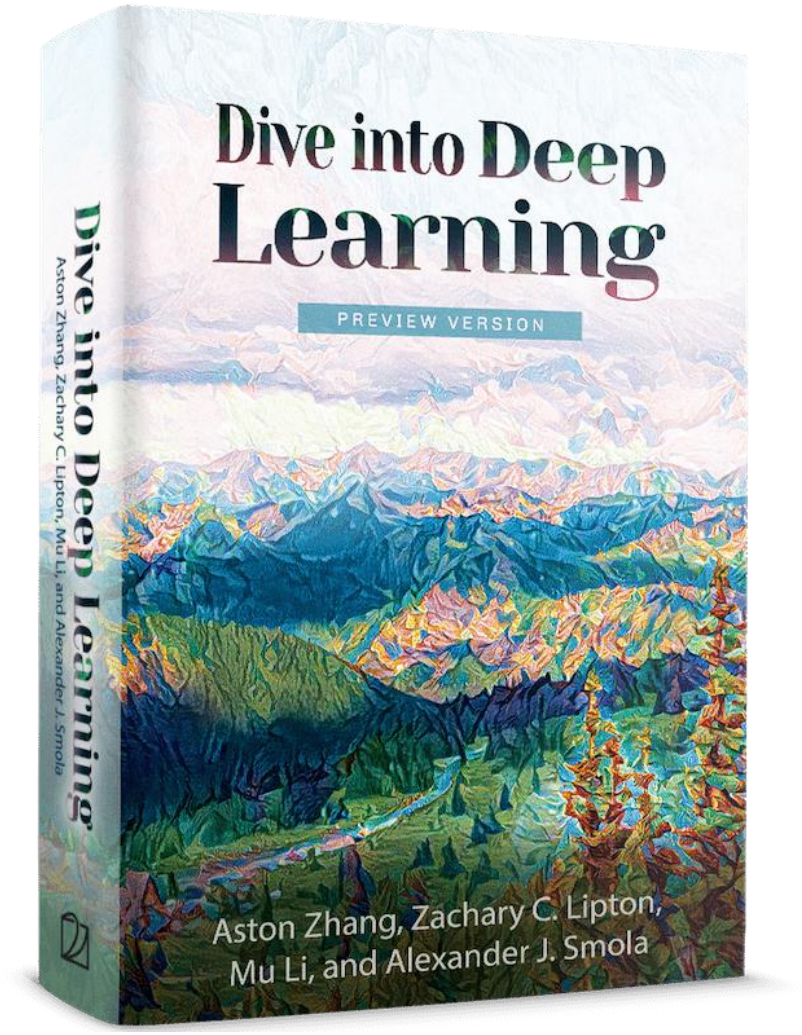
Reference book



- Available online
  - http://www.deeplearningbook.org/
  - Also in Chinese
    - Online: https://github.com/exacity/deeplearningbook-chinese
- Other references
  - Check our e3 website (NYCU e3)



- https://pytorch.org/deep-learning-with-pytorch
  - Available online



- https://d2l.ai/
  - Available online

《动手学深度学习》

第二版

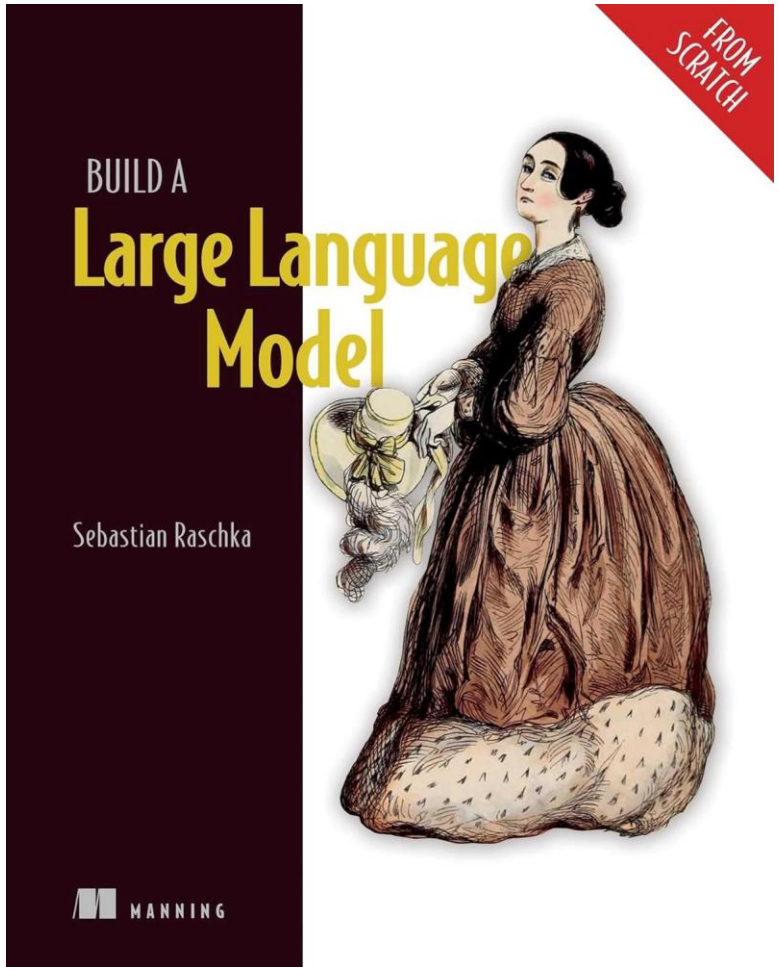
跳转第一版

面向中文读者的能运行、可讨论的深度学习教科书

含 PyTorch、NumPy/MXNet、TensorFlow 和 PaddlePaddle 实现

被全球 70 多个国家 500 多所大学用于教学

☆ Star 46,501



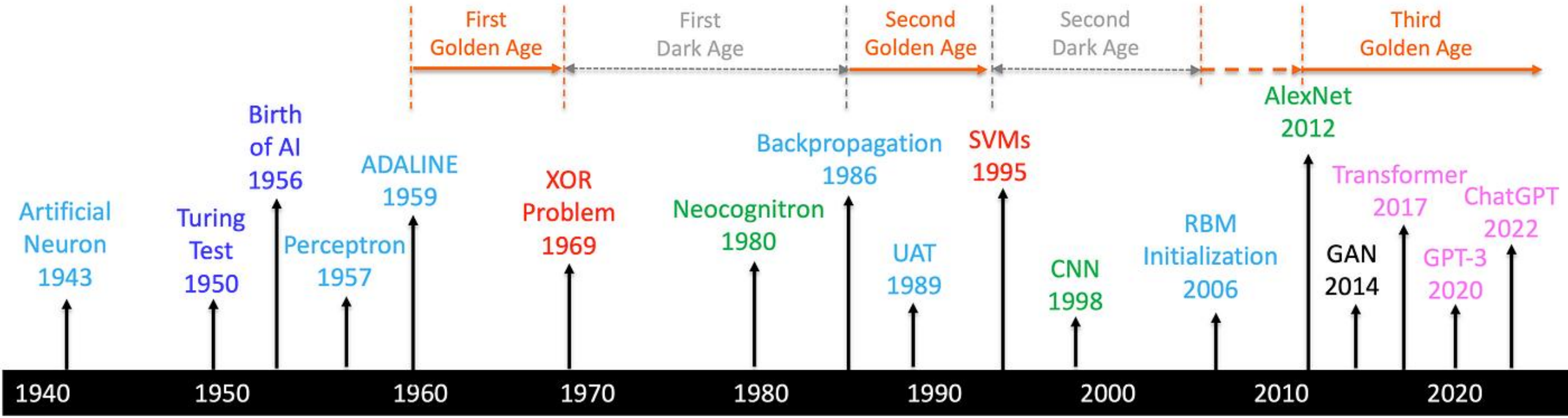
- Build a Large Language Model (From Scratch)
  - https://github.com/rasbt/LLMs-from-scratch

運算資源

- 機器學東西就是需要時間和資源
  - 作業早點開始
  - 等待以小時或天計算
- 課程能提供的資源有限，歡迎自備資源(GPU with CUDA)
- 免費運算資源
  - Google colab

Lecture 0: Introduction

A Brief History of AI with Deep Learning




McCulloch-Pitts


Rosenblatt


Widrow-Hoff


Minsky-Papert

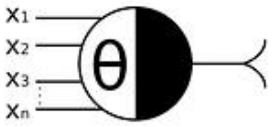

Rumelhart, Hinton et al.

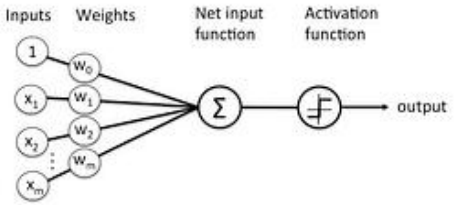

LeCun

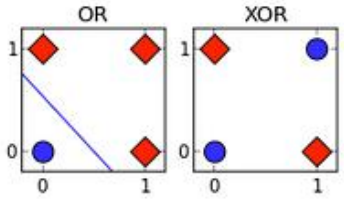

Hinton-Ruslan

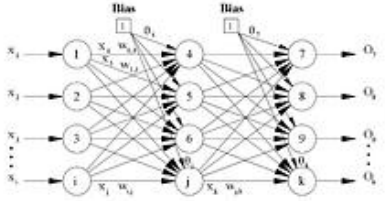

Krizhevsky et al.

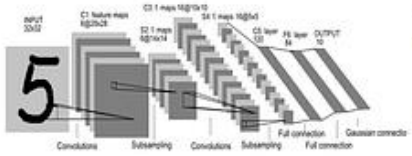

Vaswani

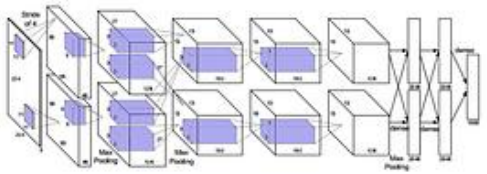




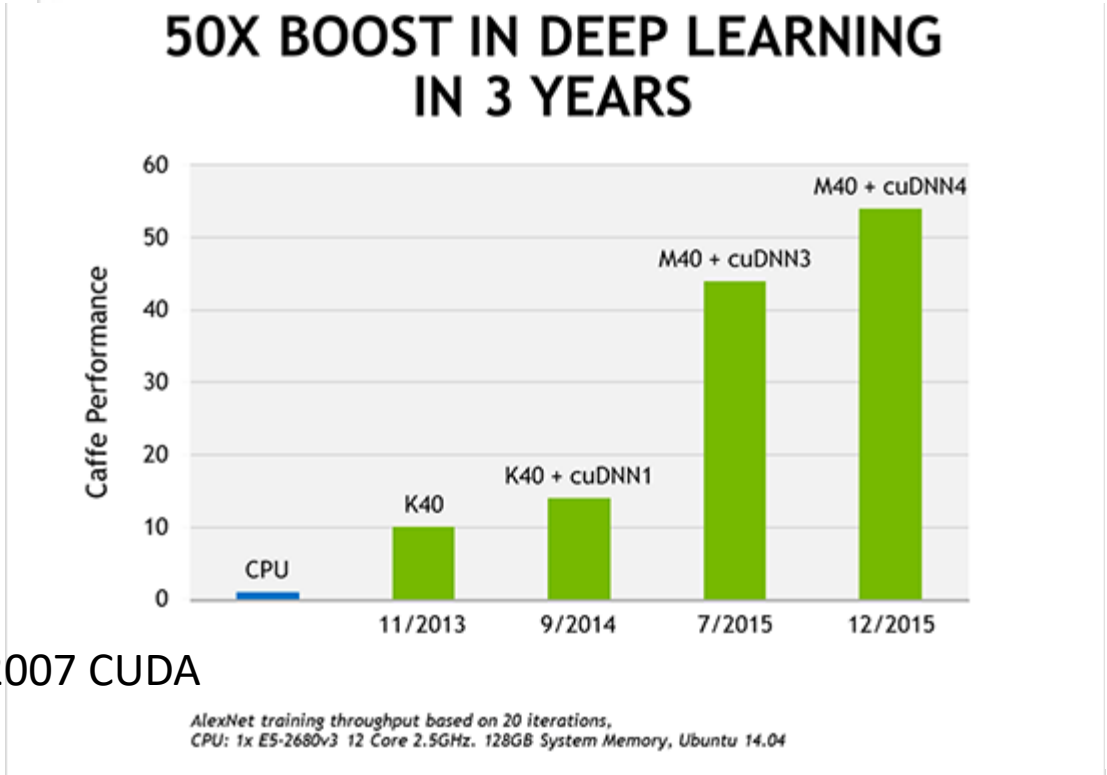




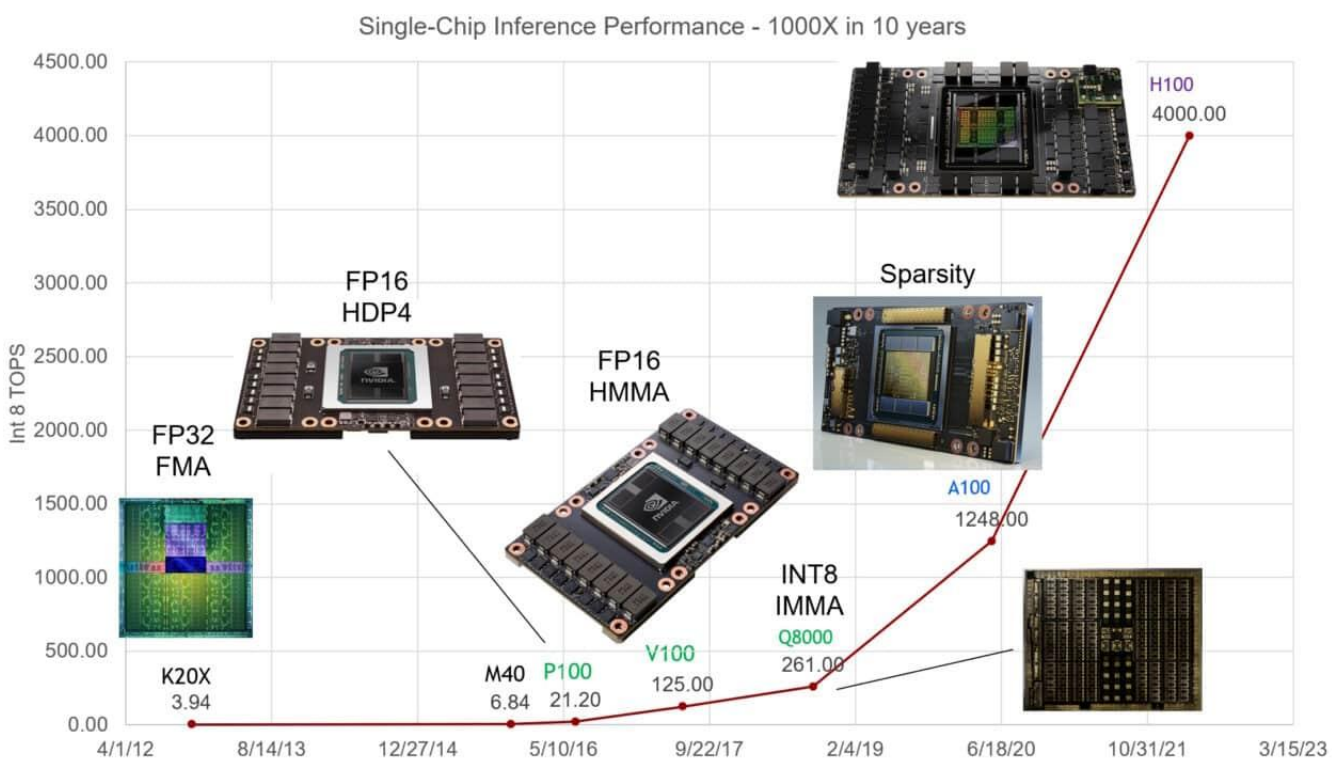




核心「資料 × 算力 × 演算法」



2007 CUDA



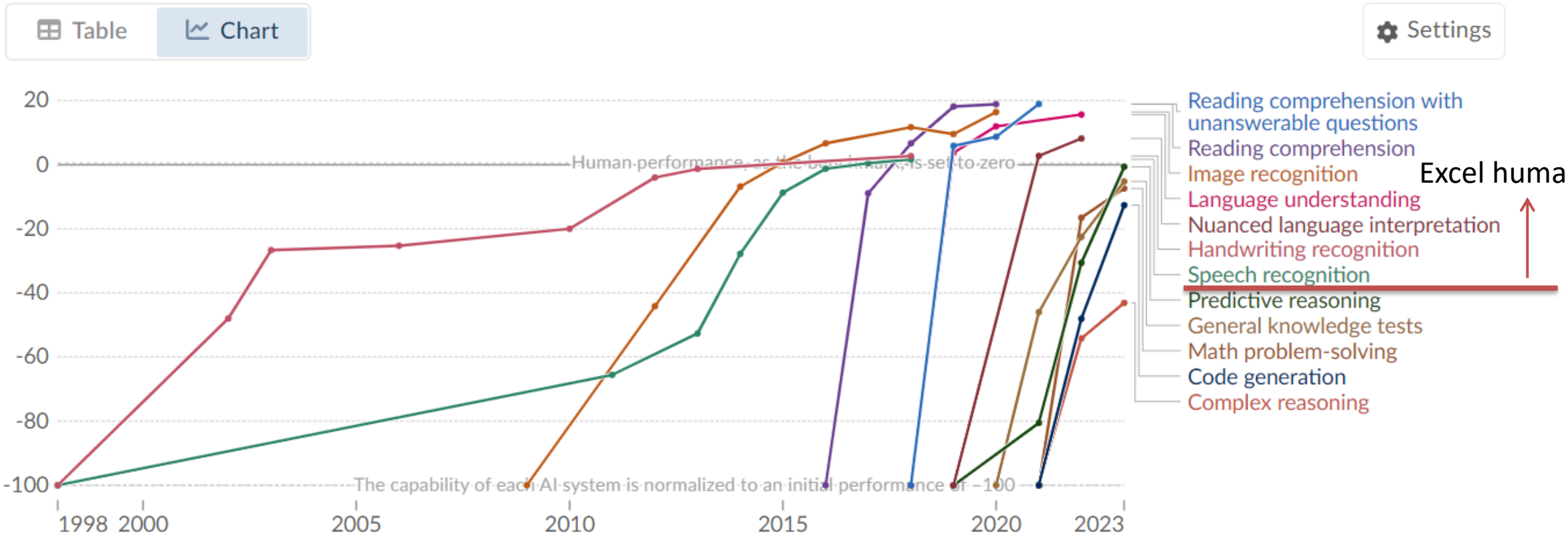
https://blogs.nvidia.com/blog/accelerating-ai-artificial-intelligence-gpus/

https://blogs.nvidia.com/blog/why-gpus-are-great-for-ai/

Test scores of AI systems on various capabilities relative to human performance

Our World in Data

Within each domain, the initial performance of the AI is set to −100. Human performance is used as a baseline, set to zero. When the AI's performance crosses the zero line, it scored more points than humans.



Data source: Kiela et al. (2023) – Learn more about this data

Note: For each capability, the first year always shows a baseline of −100, even if better performance was recorded later that year.

OurWorldInData.org/artificial-intelligence | CC BY

https://ourworldindata.org/brief-history-of-ai

Timeline of images generated by artificial intelligence

These people don't exist. All images were generated by artificial intelligence.

Our World in Data

2014

Goodfellow et al. (2014) – Generative Adversarial Networks

2015

Radford, Metz, and Chintala (2015) – Unsupervised Representation Learning with Deep Convolutional GANs

2016

Liu and Tuzel (2016) – Coupled GANs

2017

Karras et al. (2017) – Progressive Growing of GANs for Improved Quality, Stability, and Variation

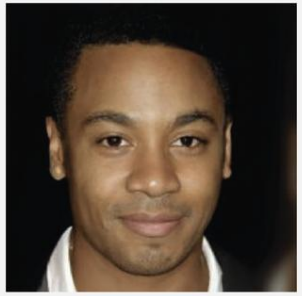
2018

Karras, Laine, and Aila (2018) – A Style-Based Generator Architecture for Generative Adversarial Networks

2019

Karras et al. (2019) – Analyzing and Improving the Image Quality of StyleGAN

2020



Ho, Jain, & Abbeel (2020) – Denoising Diffusion Probabilistic Models

2021

Image generated with the prompt: *"a couple of people are sitting on a wood bench"*

Ramesh et al. (2021) – Zero-Shot Text-to-Image Generation (OpenAI's DALL-E 1)

2022

Image generated with the prompt: *"A Pomeranian is sitting on the King's throne wearing a crown. Two tiger soldiers are standing next to the throne."*

Saharia et al. (2022) – Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding (Google's Imagen)

Explaining a Joke

I will explain these jokes:
(1) The problem with kleptomaniacs is that they always take things literally.
Explanation: This joke is wordplay. Someone who "takes things literally" is someone who doesn't fully understand social cues and context, which is a negative trait. But the definition of kleptomania is someone who literally takes things.
(2) Always borrow money from a pessimist. They'll never expect it back.
Explanation: Most people expect you to pay them back when you borrow money, however a pessimist is someone who always assumes the worst, so if you borrow money from them, they will expect that you won't pay them back anyways.

Input: I was supposed to start writing the paper at 5:00 PM. But then I started playing with this cool new language model for 10 minutes. 10 minutes later, it's suddenly 9:30 PM!
Model Output: This joke is about how time flies when you're having fun. The person was supposed to start writing a paper at 5:00 PM, but they started playing with a language model instead. They thought they were only playing with it for 10 minutes, but it was actually 4.5 hours.

Input: I think I understand why me and my partner are such a good match. I get excited when I smell food and become super alert when the doorbell rings. My girlfriend Shelley becomes ecstatic when she sees a person who she just saw 5 hours ago.
Model Output: This joke is about dogs. Dogs get excited when they smell food and when the doorbell rings. The girlfriend is excited to see someone she just saw 5 hours ago, which is a trait of dogs.

Input: I was going to fly to visit my family on April 6th. My mom said "Oh great, your step dad's poetry reading is that night!" So now I'm flying in on April 7th.
Model Output: The joke is that the speaker's mother is trying to get them to go to their step dad's poetry reading, but the speaker doesn't want to go, so they are changing their flight to the day after the poetry reading.

Input: I tried 10,000 random restarts of my neural network, but I was accused of overfitting. I guess no good seed goes unpunished.
Model Output: This joke is a pun. A neural network is a computer program that can learn from data. A "seed" is a number that is used to initialize a random number generator. A "good seed" is a number that produces a good random number generator. The phrase "no good deed goes unpunished" means that if you do something good, you will be punished for it.

Input: Reading "War and Peace" doesn't scare me. I once had to review the CL for my intern's whole project on his final day!
Model Output: 'War and Peace' is a very long book, so reading it is a daunting task. However, the speaker is saying they had to review the code for their intern's project on their final day, which is even more daunting.

Input: What's the difference between a zebra and an umbrella? One is a striped animal related to horses, another is a device you use to stop the rain from falling on you.
Model Output: This joke is an anti-joke. The joke is that the answer is obvious, and the joke is that you were expecting a funny answer.

過去幾十年來，用於訓練人工智慧的運算量持續呈指數級增長

訓練計算以對數刻度繪製，因此從每條網格線到下一條網格線，它顯示出 100 倍的增長。從長遠來看，這一趨勢顯示出持續成長。在最初的 60 年裡，訓練計算量按照摩爾定律增長，大約每 20 個月翻倍。大約自 2010 年以來，這種指數增長進一步加快，僅用 6 個月就翻了一番。這是一個驚人的快速成長速度。

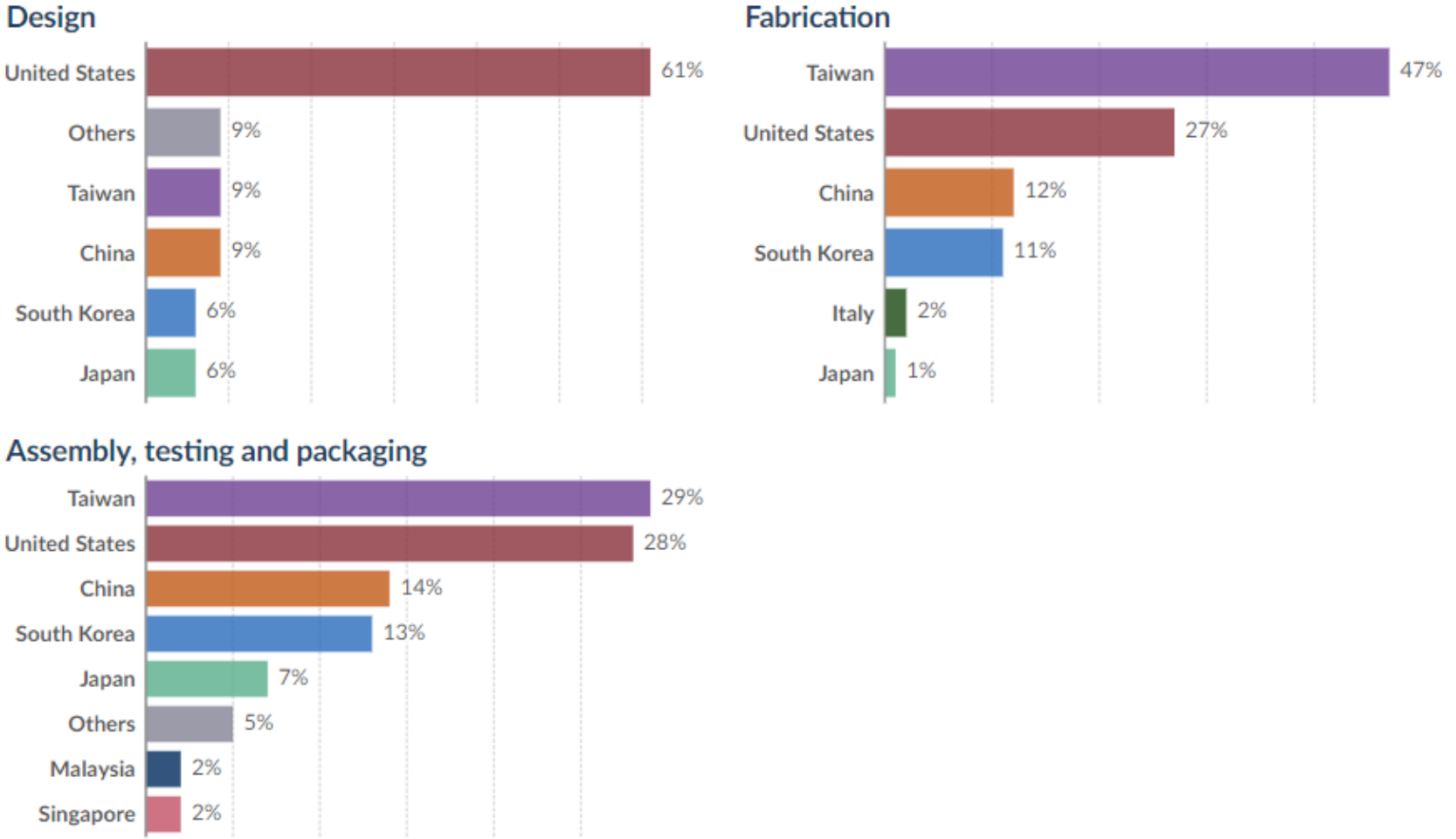
Market share for logic chip production, by manufacturing stage, 2021

Collective market share of all firms headquartered in a country. Logic chips, such as CPUs and GPUs, are the fundamental information-processing units of computers and other electronic devices.

Our World in Data

Table | Chart

Settings



Data source: Center for Security and Emerging Technology (2022) – Learn more about this data

OurWorldInData.org/artificial-intelligence | CC BY

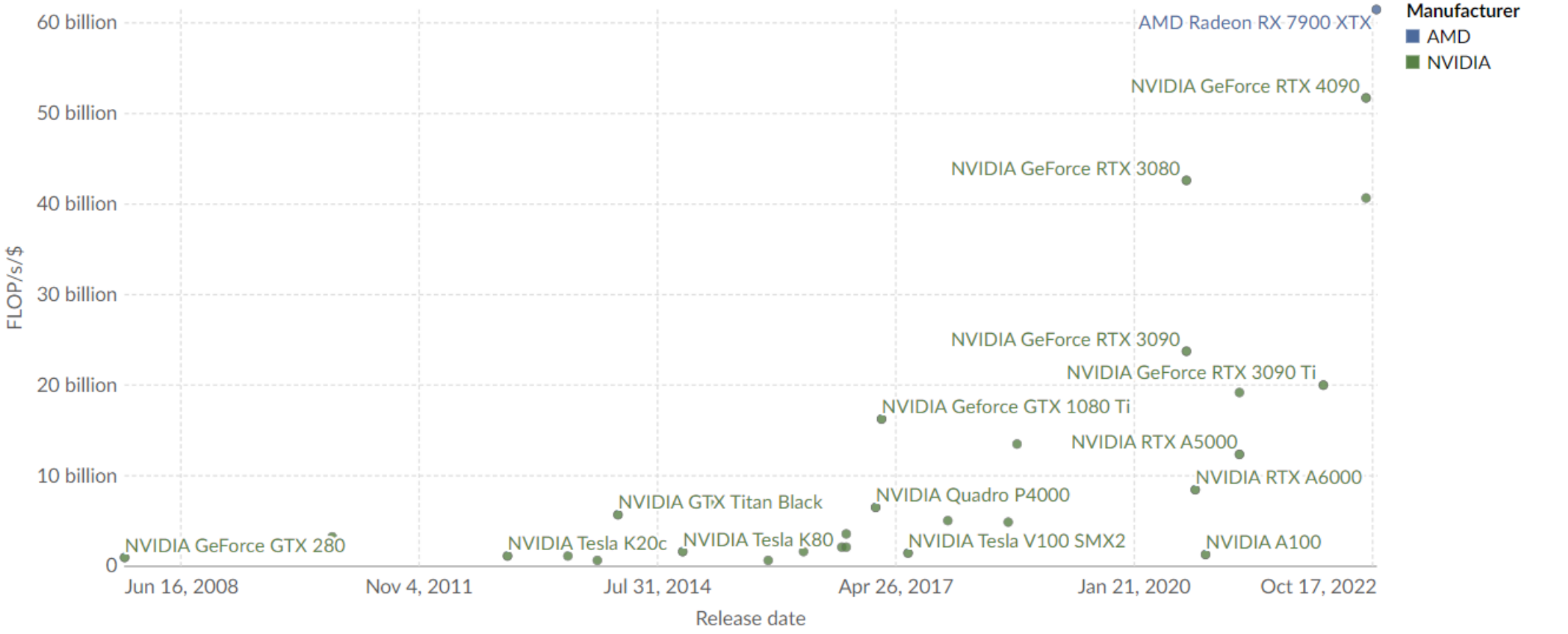
每美元的 GPU 運算效能

圖形處理單元 (GPU) 是人工智慧系統的主要運算硬體。 GPU效能以浮點顯示每美元每秒的操作次數 (FLOP/s)，根據通貨膨脹進行調整。

Our World in Data

桌子 圖表

設定



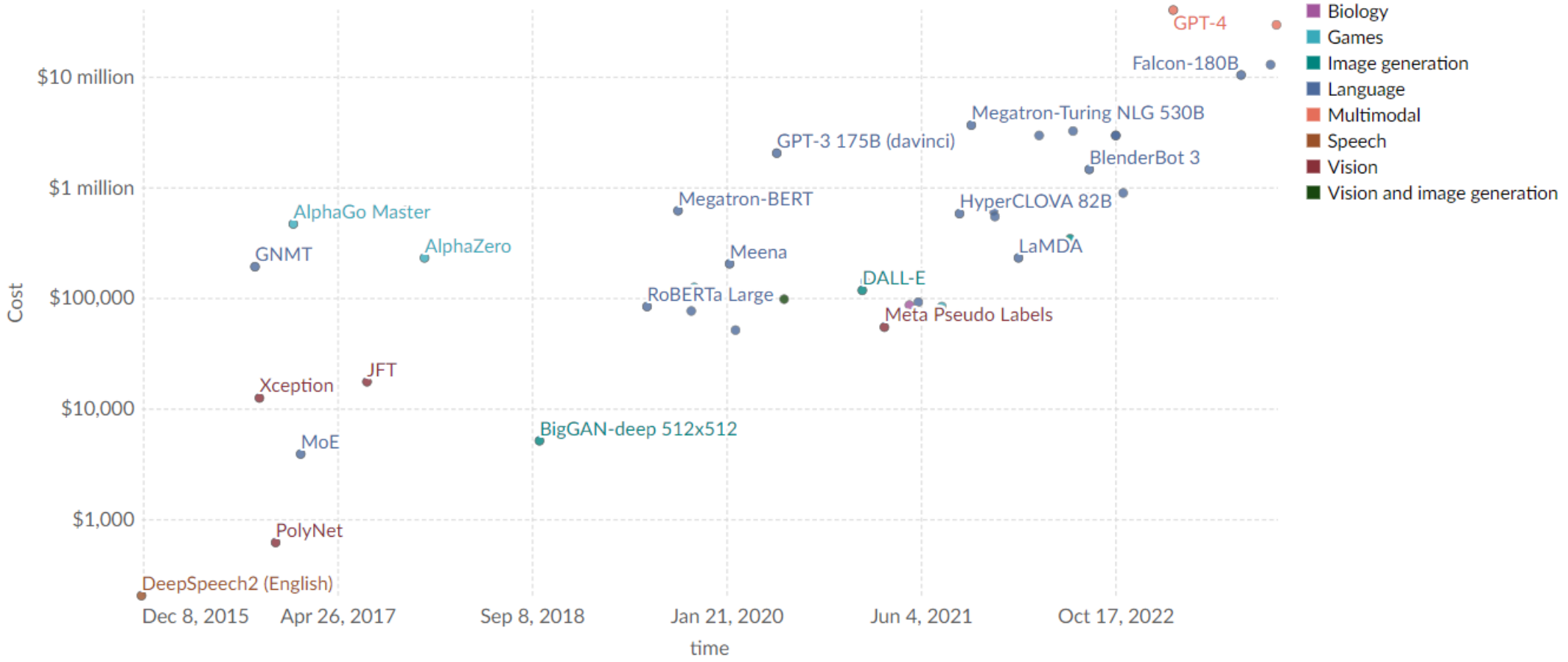
播放縮時攝影 2008 年 6 月 16 日 2022 年 11 月 3 日

訓練著名人工智慧系統的硬體和能源成本

該數據以美元表示，並根據通貨膨脹進行調整。

Our World in Data

桌子 圖表 設定



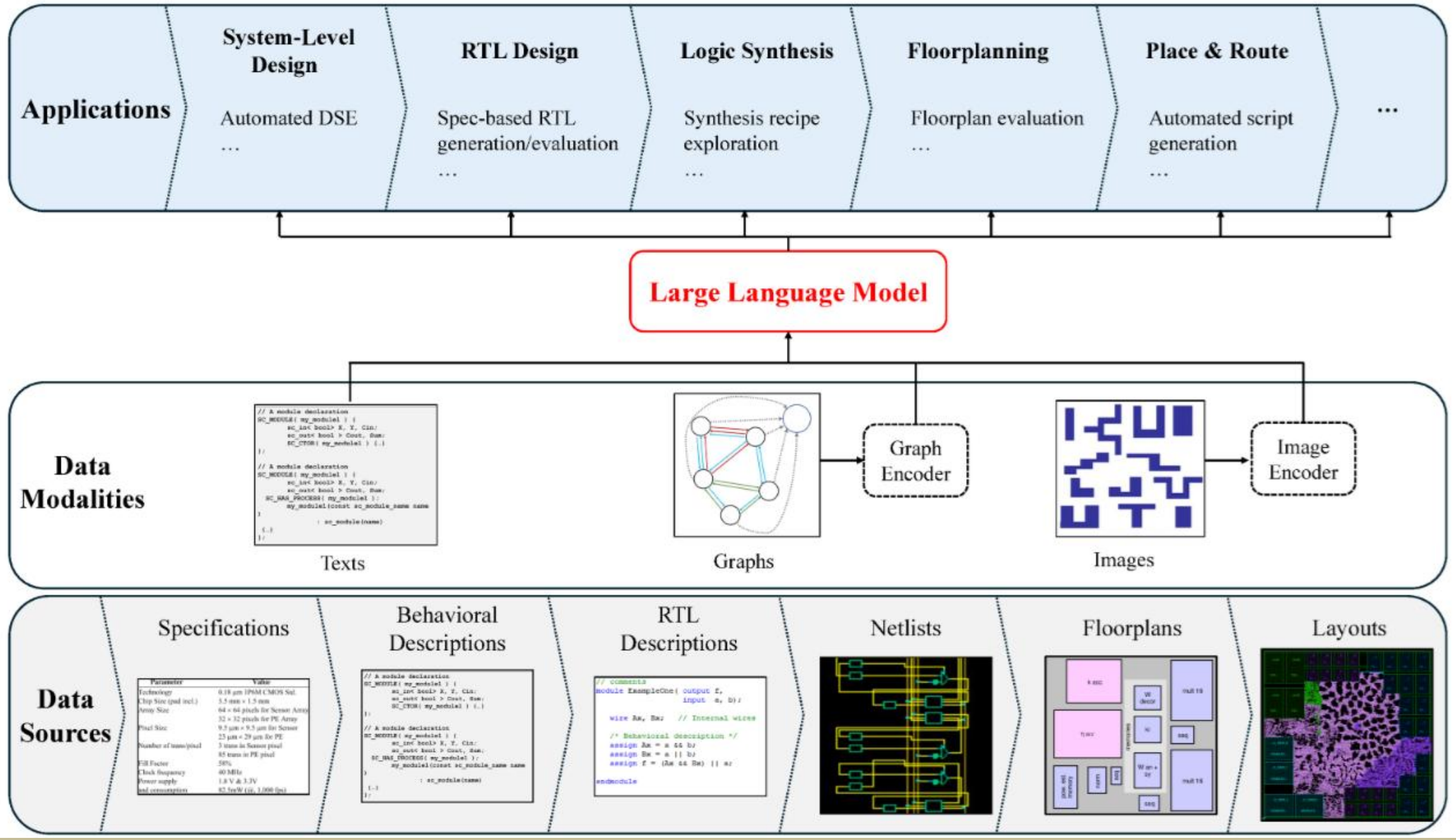
播放縮時攝影 2015 年 12 月 8 日 2023 年 12 月 6 日

Yann LeCun
@ylecun

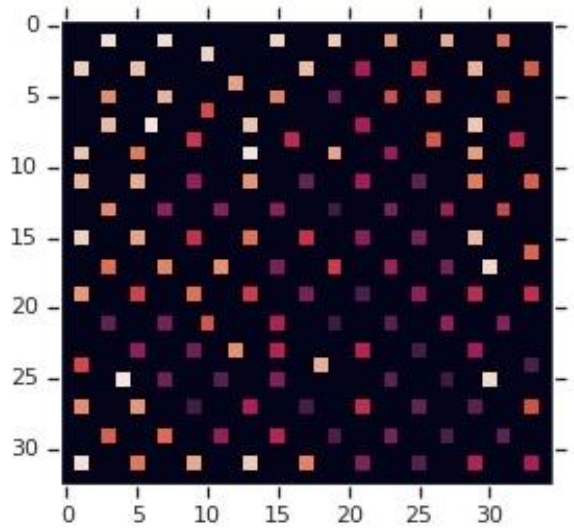
Let's start with Artificial Rat-level Intelligence (ARI), then move on to Artificial Cat-level Intelligence (ACI), and so on to Artificial Human-level Intelligence (AHI).

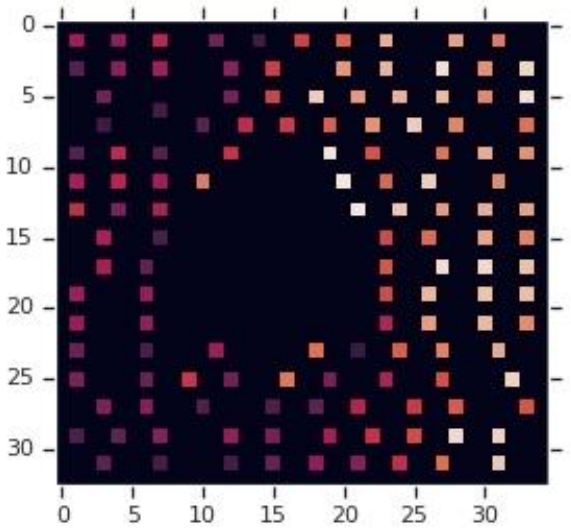
AGI does not exist, other than as Artificial {Gerbil,Goose,Giraffe,Gibbon,Gorilla} -level Intelligence....

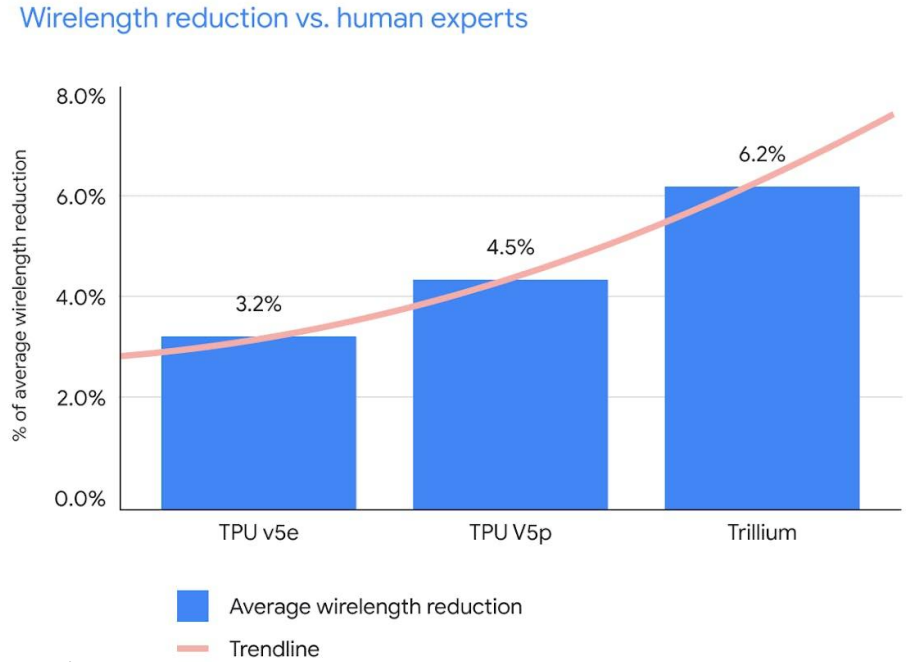
LLM for Digital Design



Using AI to design Google's AI accelerator chips







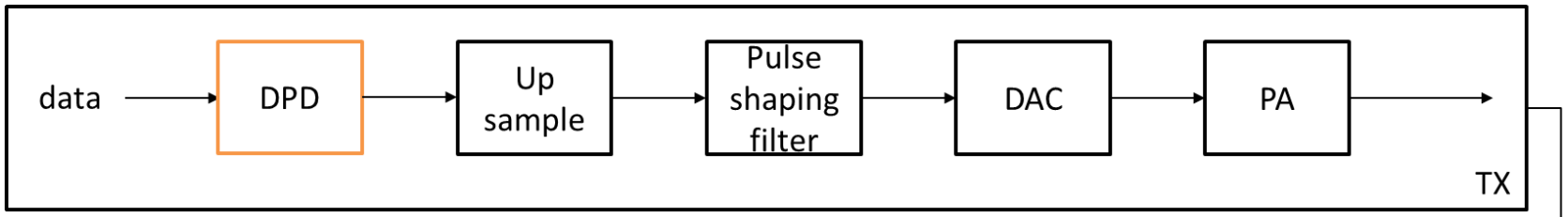
Google DeepMind 就「AlphaChip」，而且早已幫 Google 實際生產出多款 TPU，只要短短數小時，就可以完成傳統完成人類設計師需要花上幾個禮拜，甚至幾個月的晶片設計工作。

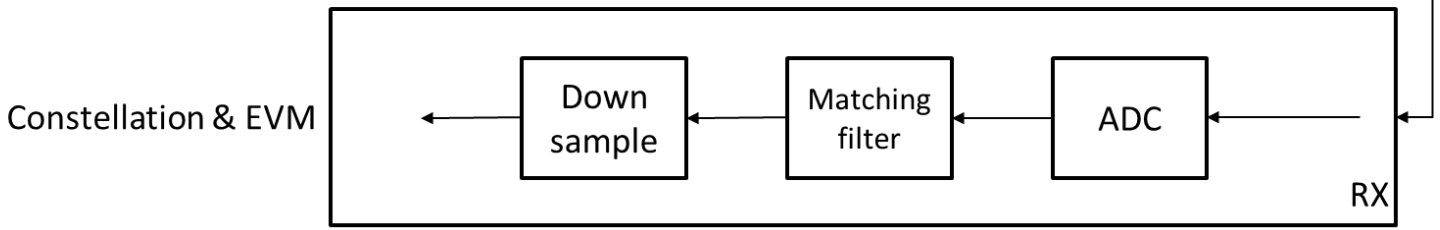
AlphaChip 的核心理念，是將晶片佈局設計視為一種「對弈」（非常接近 AlphaGo 的做法），並利用強化學習算法來訓練 AI。簡單來說，AlphaChip 會從一個空白的晶片網格出發，玩拼圖一樣逐步放置電路元件，並根據佈局的優劣獲得相應的獎勵。通過不斷地訓練和學習，AlphaChip 最終掌握了設計高效晶片佈局的訣竅。

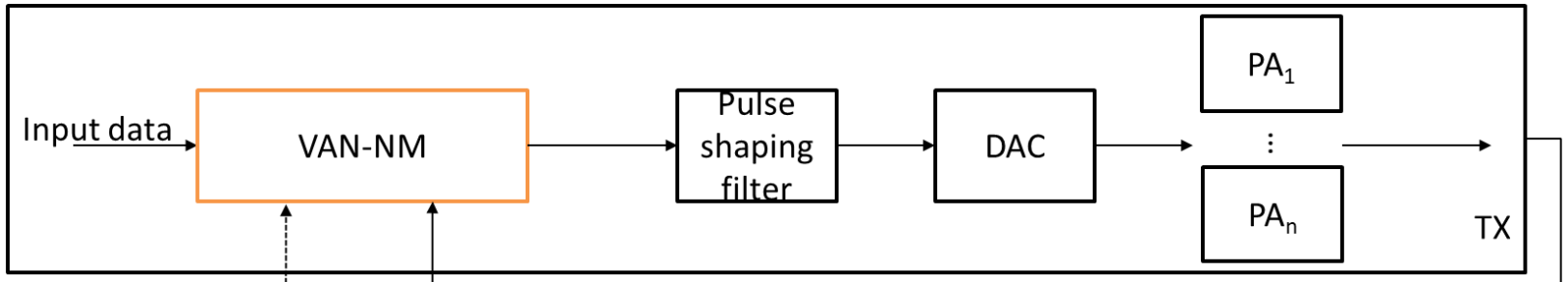
https://deepmind.google/discover/blog/how-alphachip-transformed-computer-chip-design/

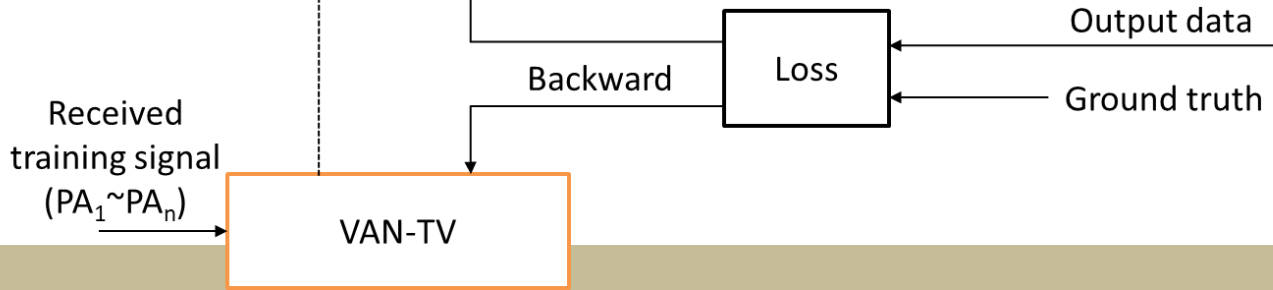
AI for Analog

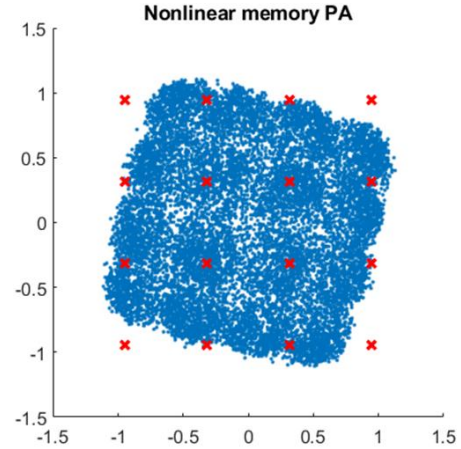
Baseband Transceiver Architectures => Non-ideal power amplifier

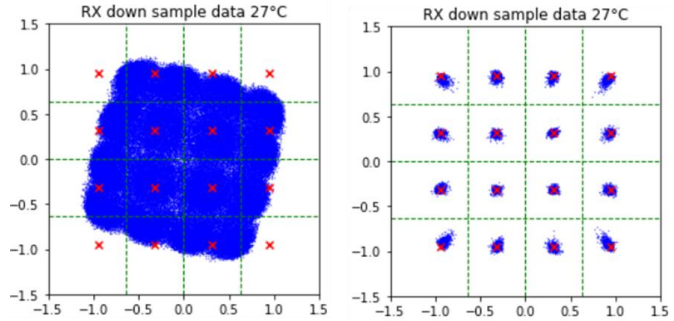






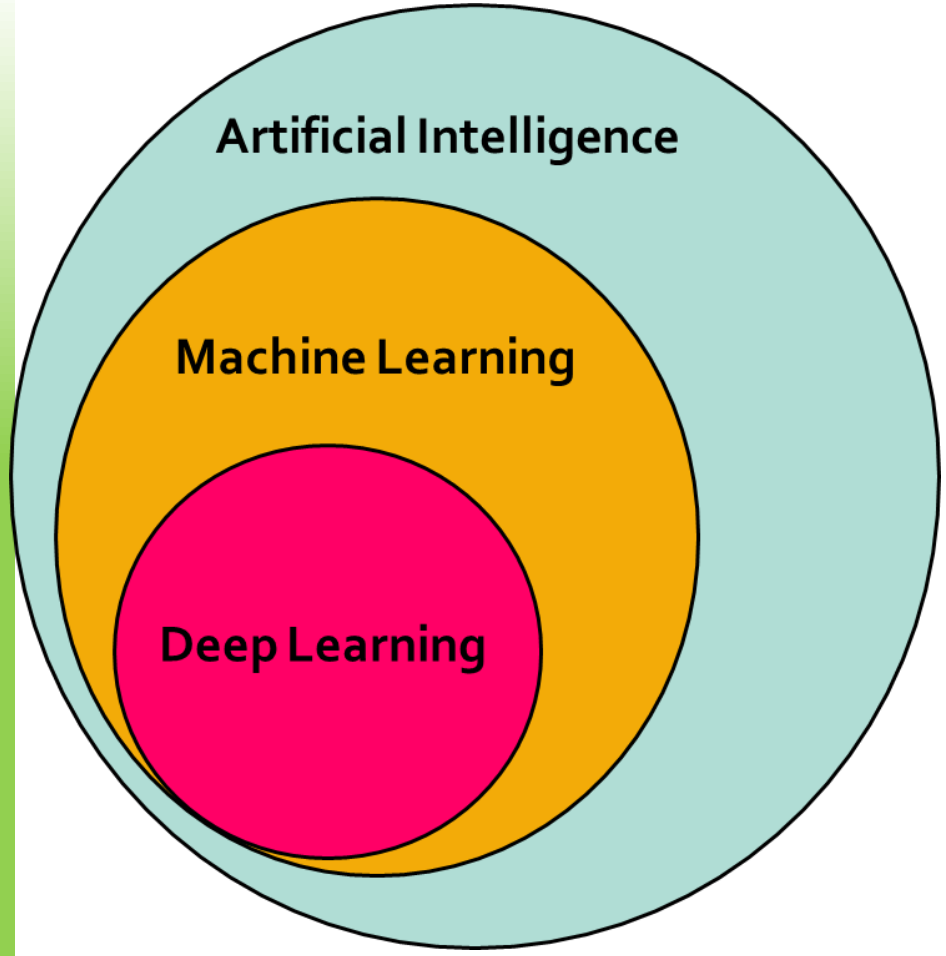






WHY DEEP LEARNING MODELS

AI vs Machine Learning vs Deep Learning



- AI: 模擬人類智慧
  - 結果有智慧就算
  - 一個擁有非常詳盡的 rule-based 系統也可以是 AI
- Machine learning是達成 AI 的一種方法
  - 從資料當中學習出 rules (不直接寫程式)
  - 找到一個夠好的 function 能解決特定的問題
- Deep learning 是machine learning的一種
  - 從feature engineering 走向architecture engineering
  - 不再人工萃取特徵
  - 深層網路萃取更抽象特徵

Deep Learning v.s. Feature Engineering

Raw data: pixel grid		
Better features: clock hands' coordinates	{x1: 0.7, y1: 0.7} {x2: 0.5, y2: 0.0}	{x1: 0.0, y2: 1.0} {x2: -0.38, 2: 0.32}
Even better features: angles of clock hands	theta1: 45 theta2: 0	theta1: 90 theta2: 140

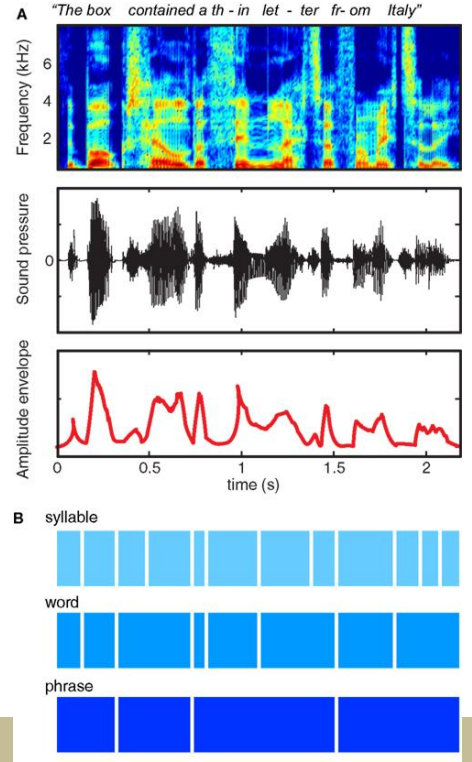
- 讓機器看時鐘報時
- 直接看圖
  - 要用CNN才行
  - 需要大量資料
- 放點工人智慧
  - 用指針座標
  - 簡單的ML就可以
  - 少量資料就可以
- 更多工人智慧
  - 用指針角度(像人看時鐘一樣)
  - 連ML都不用，查表就可以
  - 資料最少

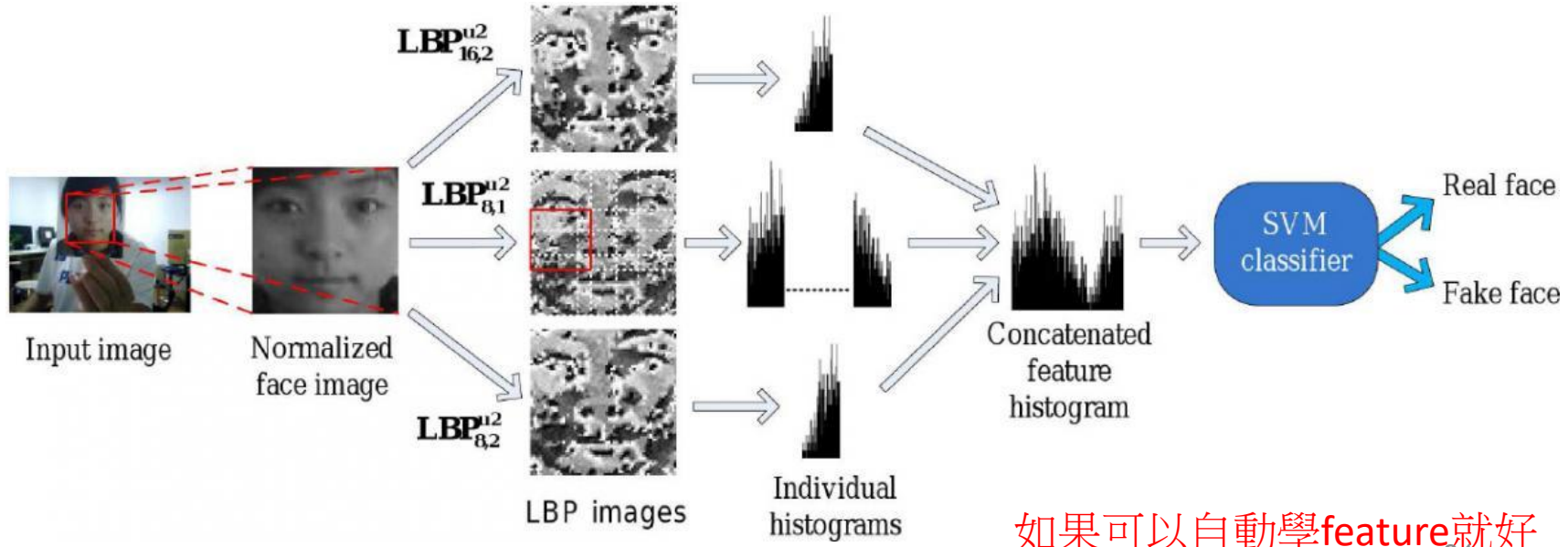
對DL來講，好的特徵可以幫助你用較少資源與資料，
反過來，若你的資料資源很少，你會需要比較好的特徵(aka.更多工人智慧)

Source: Deep learning with python

Before Deep learning

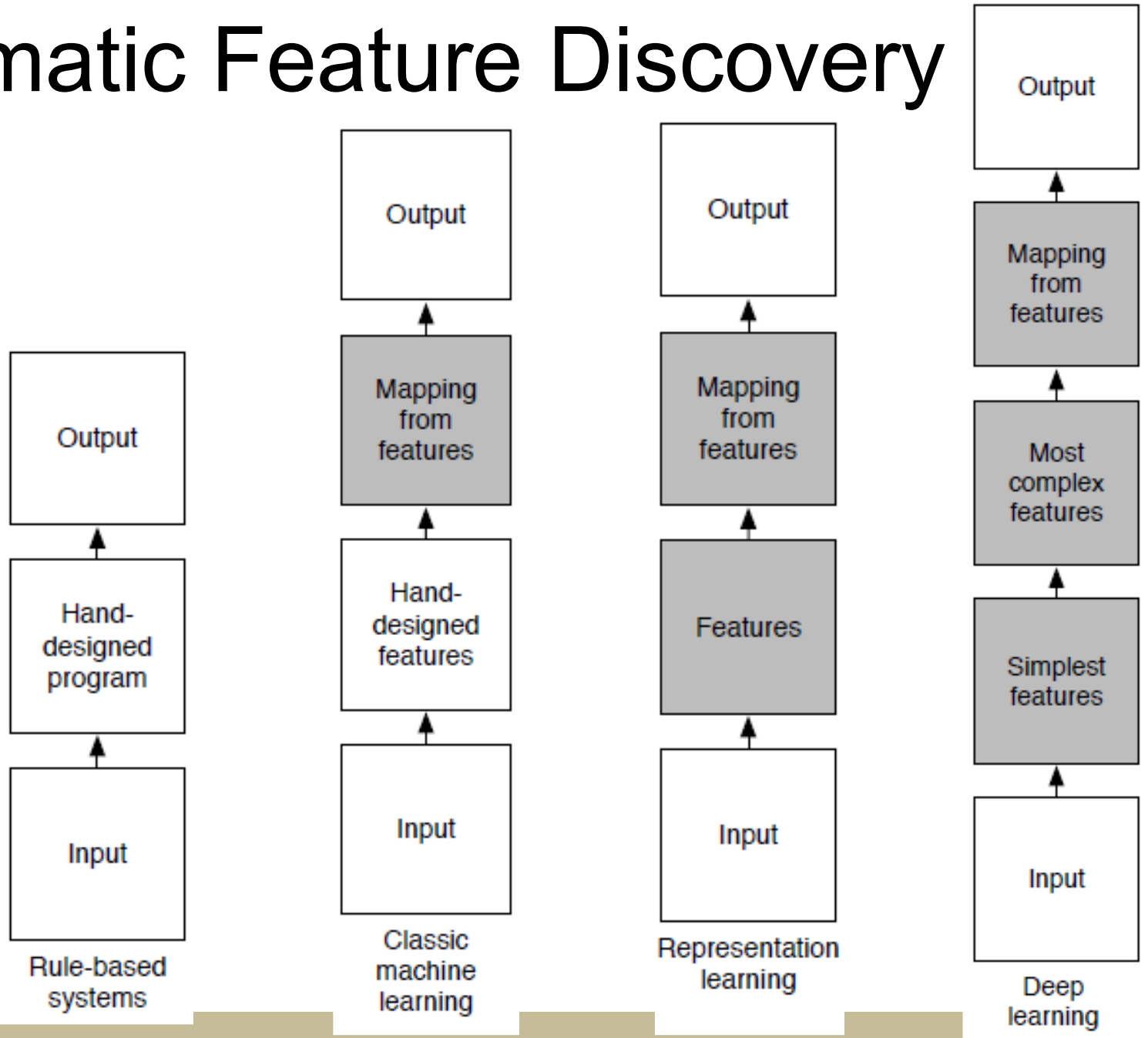
- Use conventional ML, e.g. SVM, you have to define features first
- 1-line summary of ML: $y = \mathrm{sgn}(W^T x + b)$
  - SVM can learn very effective weights w
    - If you use the right representation x ,aka. features





如果可以自動學feature就好

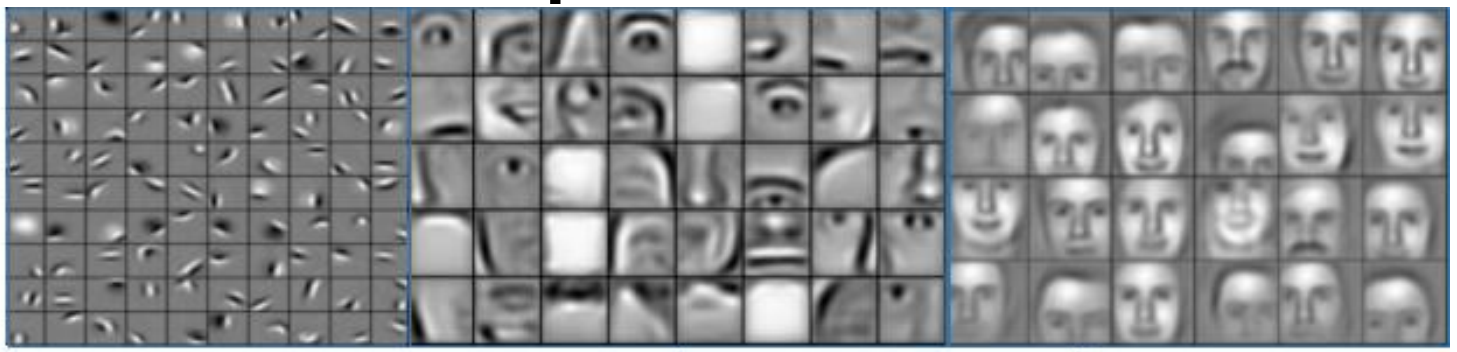
DL: Automatic Feature Discovery

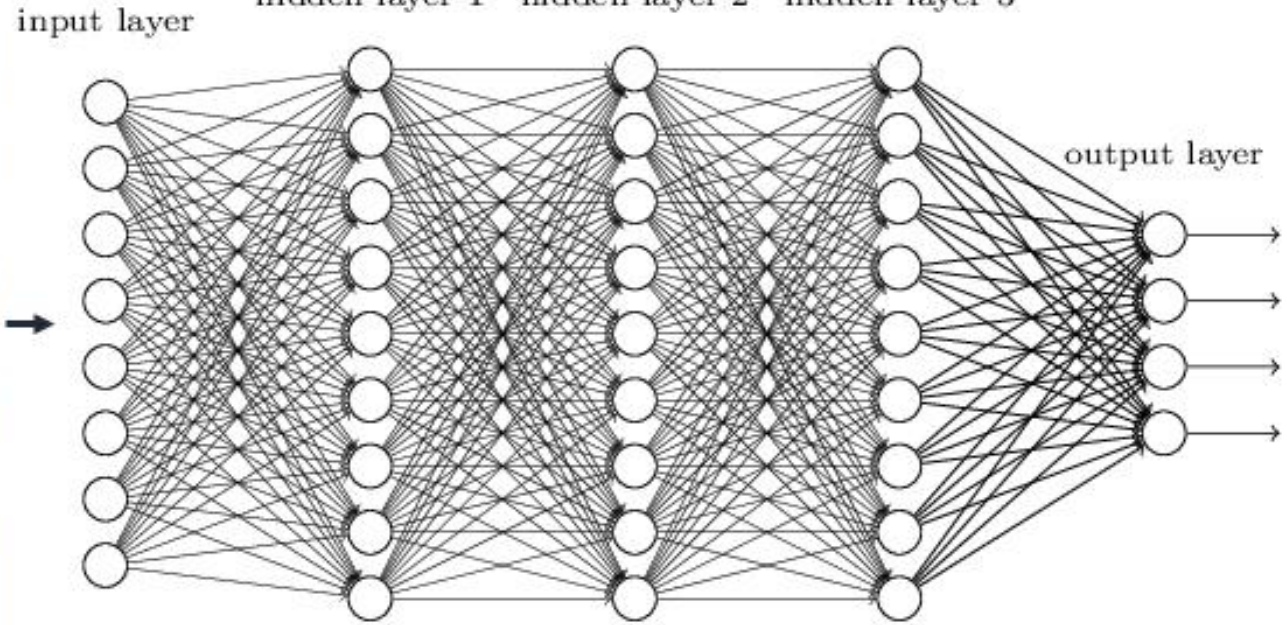


DL到底學了什麼?越深越好嗎?

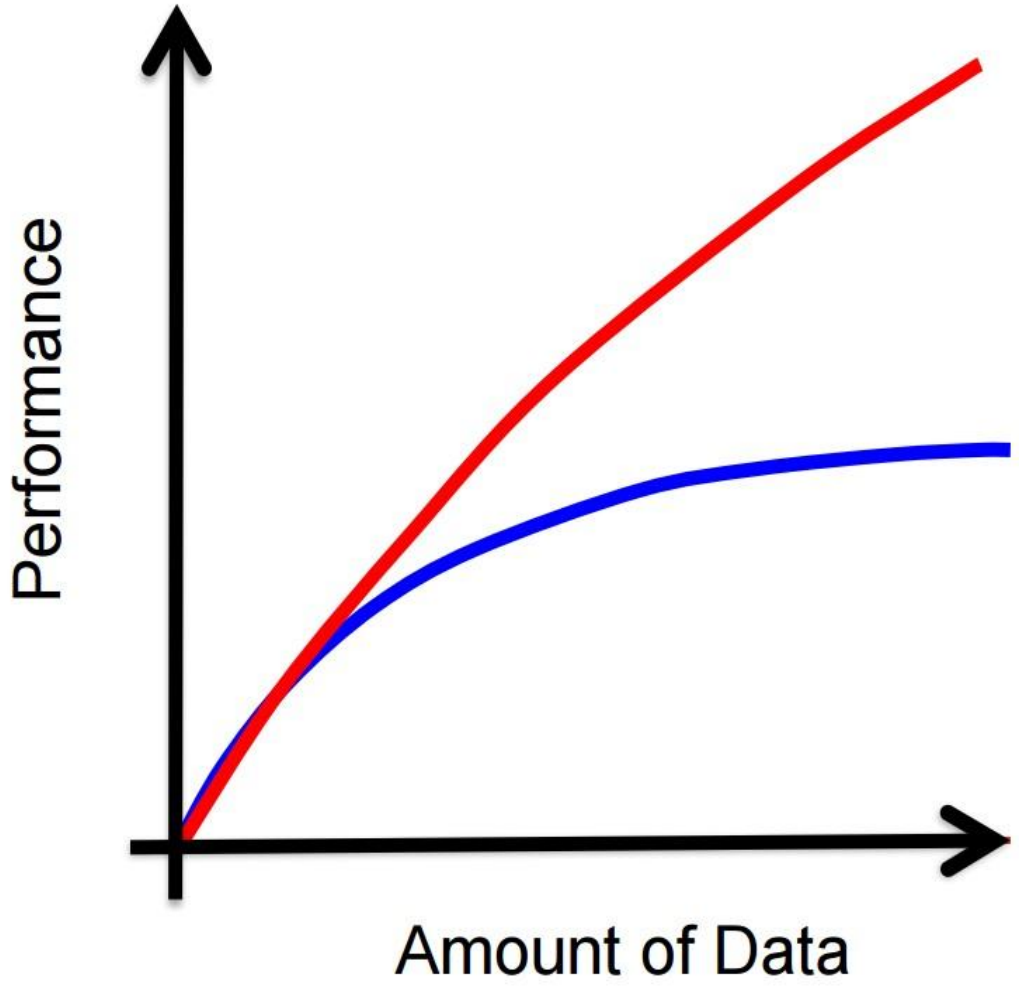
Hierarchical Feature Representations

Deep neural networks learn hierarchical feature representations





Deep Learning: **Scalable** Machine Learning

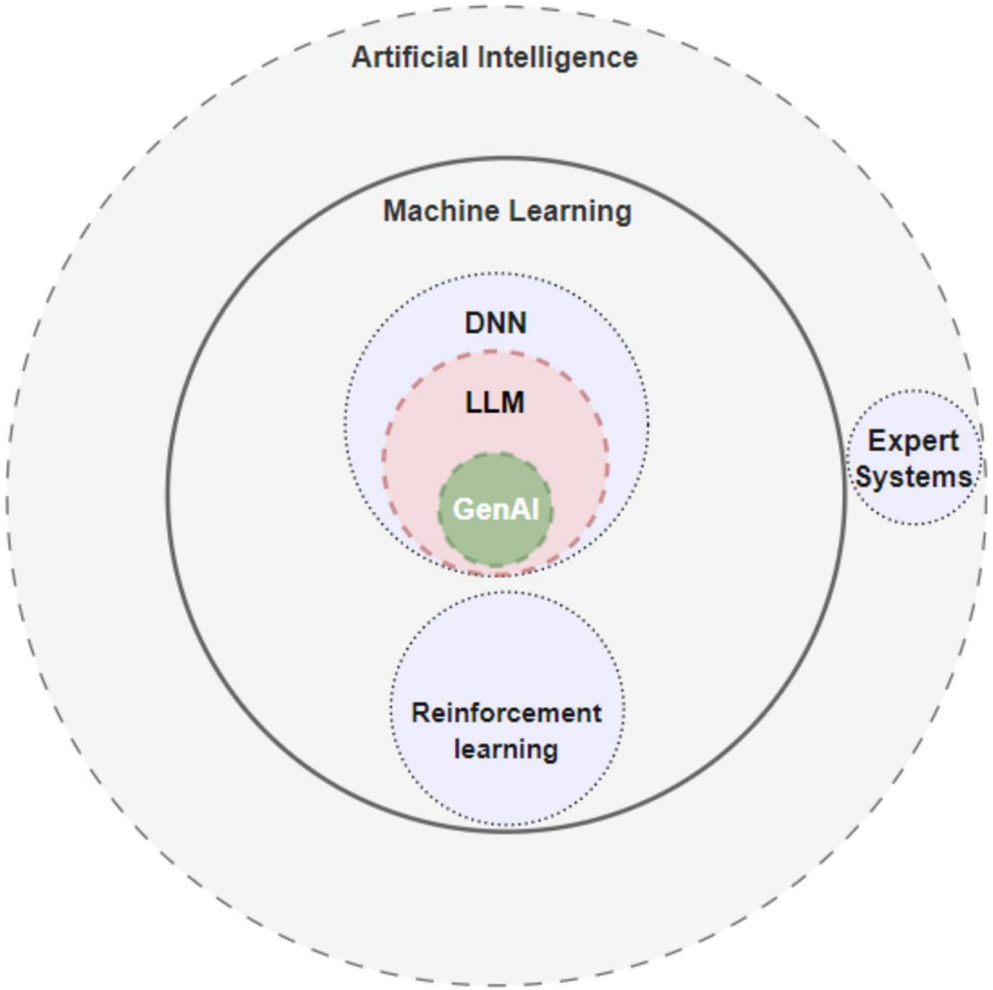


從資料自動學特徵
資料越多特徵越多，效果越好

Feature engineering
需要人工設計特徵
無法隨更多資料自動學更多特徵

Source: MIT, Deep Learning for Self-Driving Cars

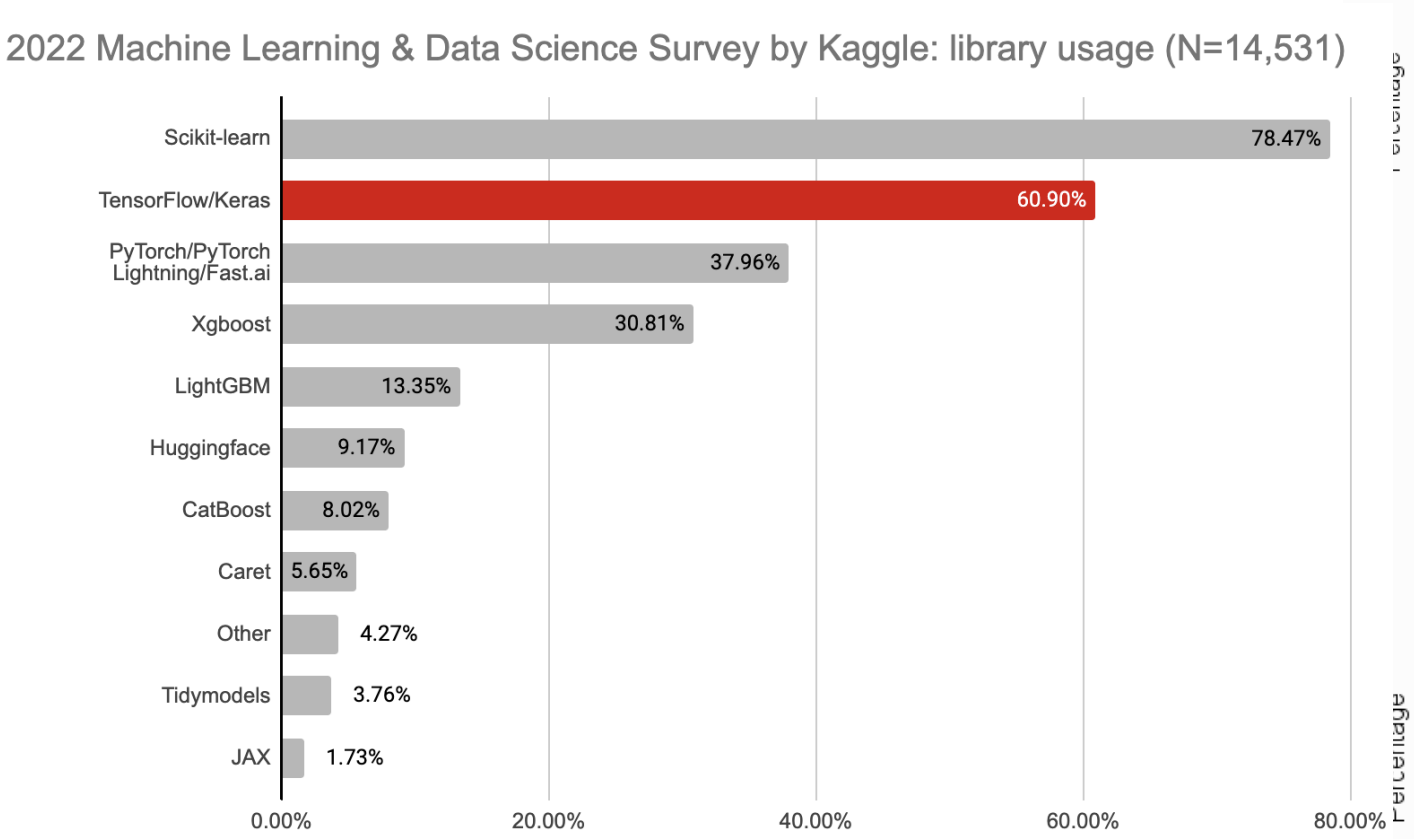
AI 的類型

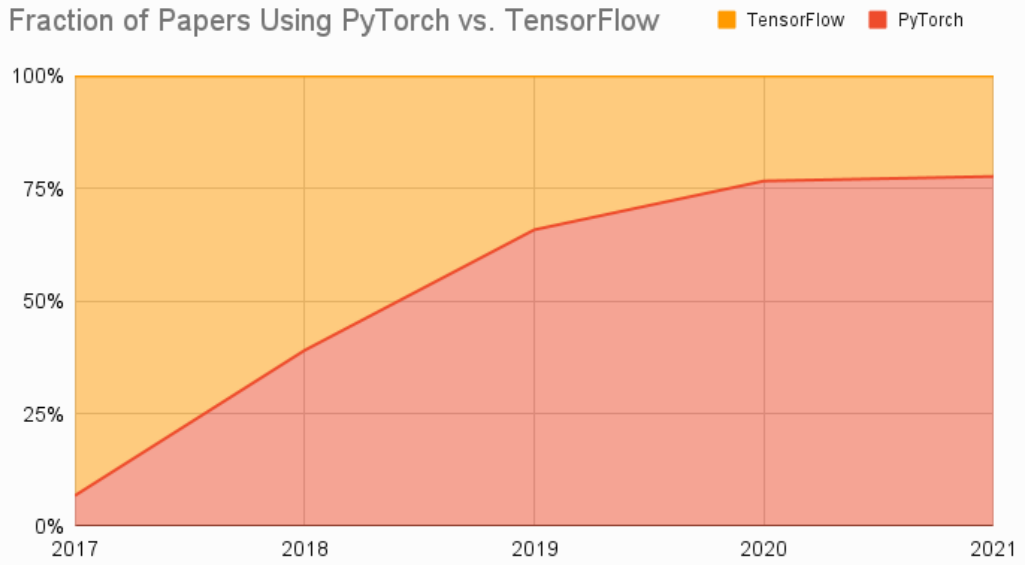


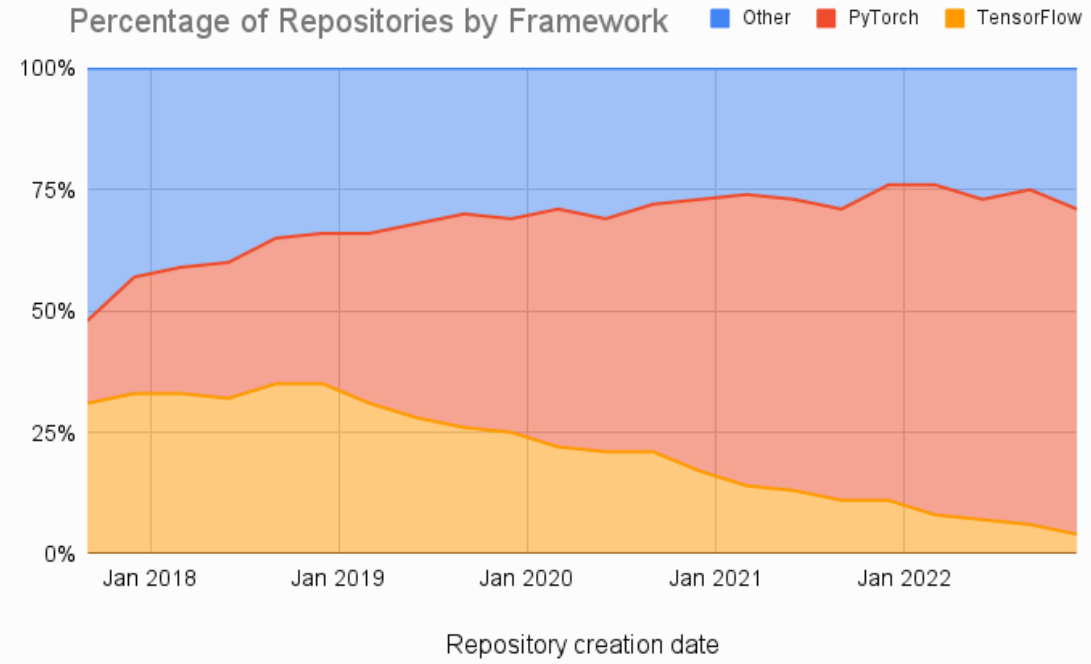
基本上既是用不同的方式
去學函數f(x)

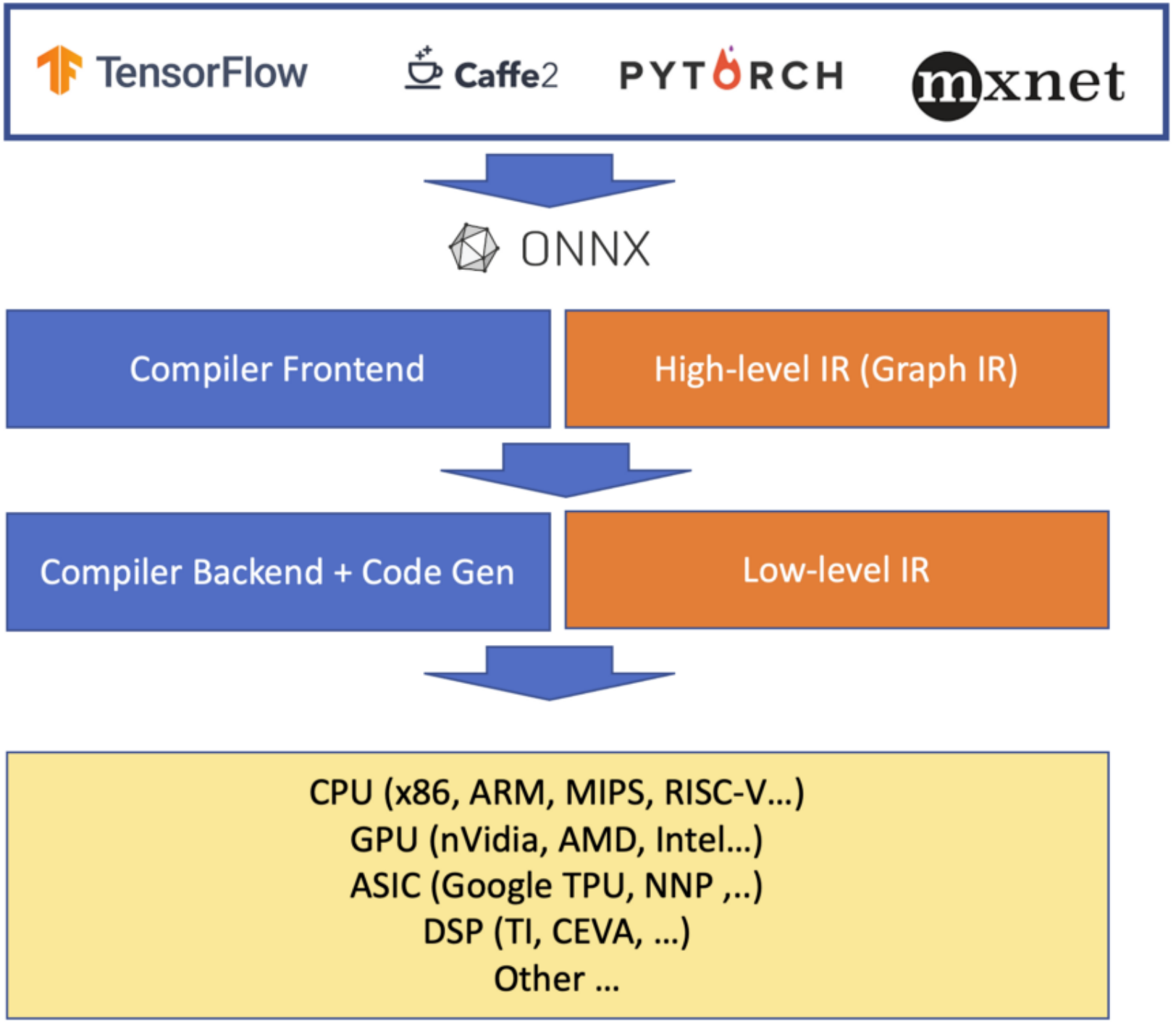
DEEP LEARNING FRAMEWORK

https://www.slideshare.net/yenlung/presentations


2022 Machine Learning & Data Science Survey by Kaggle: library usage (N=14,531)
Scikit-learn
78.47%
TensorFlow/Keras
60.90%
PyTorch/PyTorch Lightning/Fast.ai
37.96%
Xgboost
30.81%
LightGBM
13.35%
Huggingface
9.17%
CatBoost
8.02%
Caret
5.65%
Other
4.27%
Tidymodels
3.76%
JAX
1.73%
0.00%
20.00%
40.00%
60.00%
80.00%


Fraction of Papers Using PyTorch vs. TensorFlow
TensorFlow
PyTorch
100%
75%
50%
25%
0%
2017
2018
2019
2020
2021


Percentage of Repositories by Framework
Other
PyTorch
TensorFlow
100%
75%
50%
25%
0%
Jan 2018
Jan 2019
Jan 2020
Jan 2021
Jan 2022
Repository creation date



Open source:
TVM