

从科技到泡沫 踉踉跄跄的AI发展史

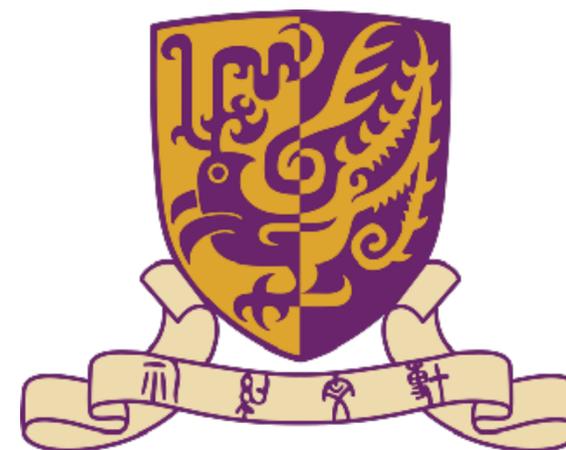
—— AI简史

刘圣超

香港中文大学 计算机科学与工程系 助理教授
波动智能实验室 (Wave Intelligence Lab) 负责人
主页: <https://chao1224.github.io/>

之前经历:

加州大学伯克利分校(UCB) 博士后
蒙特利尔算法研究中心(Mila) 博士



AI简史

1950s: AI的起源

- 1950 艾伦 图灵 (Alan Turing)提出了图灵测试
- 1956 达特茅斯研讨会提出了'人工智能' (Artificial Intelligence)这一术语, [workshop link](#)

“A test of a machine's ability to exhibit intelligent behavior equivalent to that of a human.”
— Alan Turing, *Computing Machinery and Intelligence*, 1950.

A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

J. McCarthy, Dartmouth College
M. L. Minsky, Harvard University
N. Rochester, I.B.M. Corporation
C.E. Shannon, Bell Telephone Laboratories

August 31, 1955

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. 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. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

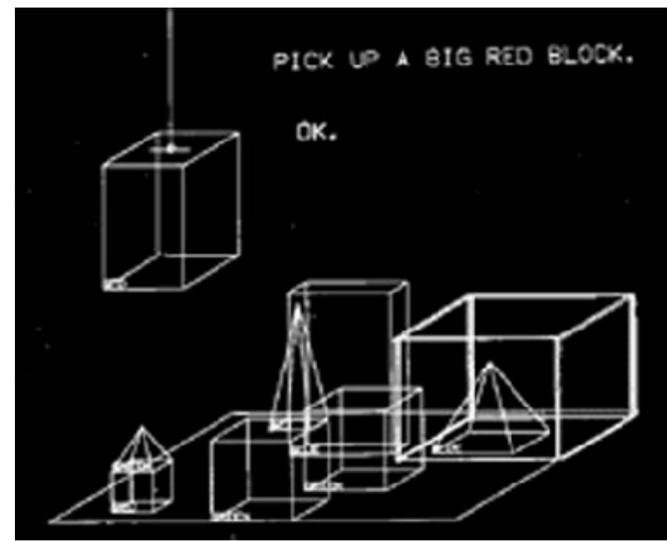
AI简史

1950s: AI的起源

- 1950 艾伦 图灵 (Alan Turing)提出了图灵测试
- 1956 达特茅斯研讨会提出了'人工智能' (Artificial Intelligence)这一术语, [workshop link](#)

1950s-1970s: 符号主义AI与基于规则的系统

- 专家系统
 - 1996 ELIZA, 一个早期的 聊天机器人模拟心理治疗师
 - 1970 SHRDLU, 虚拟积木世界



```
Welcome to
EEEEEE LL      IIII  ZZZZZZ  AAAAA
EE      LL      II    ZZ     AA  AA
EEEEEE LL      II    ZZZ    AAAAAA
EE      LL      II    ZZ     AA  AA
EEEEEE LLLLLL  IIII  ZZZZZZ  AA  AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:
```

AI简史

1980s: 机器学习(Machine Learning)的兴起

- 1986 David E. Rumelhart, Geoffrey Hinton 等人提出 **反向传播算法**
- 1989 Yann LeCun 提出 **卷积神经网络(Convolution Neural Network)**, 用于手写数字识别



AI简史

1980s: 机器学习(Machine Learning)的兴起

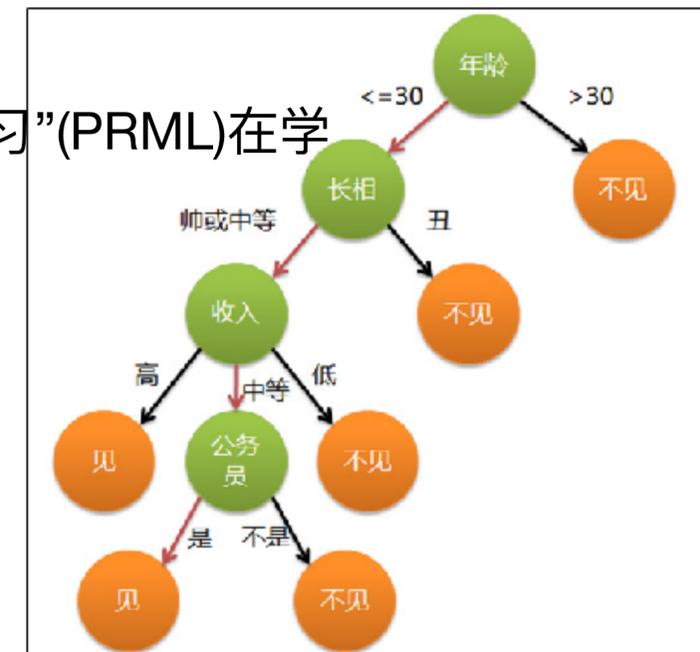
- 1986 David E. Rumelhart, Geoffrey Hinton 等人提出 **反向传播算法**
- 1989 Yann LeCun 提出 **卷积神经网络(Convolution Neural Network)**, 用于手写数字识别



Tom M. Mitchell, MSR

1990s: 统计机器学习的到来

- **决策树(Decision Trees)**、**支持向量机(SVM)**、**K近邻(KNN)**、**朴素贝叶斯(Naive Bayes)** 等方法得到广泛研究和应用
 - 李开复在语音识别应用中用的是朴素贝叶斯
- Tom M. Mitchell 的机器学习教材, “**模式识别与机器学习**”(PRML)在学术界具有重要影响
 - 有微软研究院(Microsoft Research/MSR)主导
 - 微软研究院成为机器学习的重要机构之一



AI简史

2000s - 至今: 深度学习 (Deep Learning)的称霸

- 2006 Geoffrey Hinton 提出 深度置信网络(Dep Belief Networks), 并推广“深度学习 (Deep Learning)”这一概念
- 2008 吴恩达 (Andrew Ng) 团队推动使用GPU进行深度学习计算
- 2012 AlexNet 在 ImageNet图像识别比赛 中获胜 (由 Geoffrey Hinton 团队提出)
- 2013 变分自编码器(Variational Autoencoder/VAE) 由 Diederik Kingman 和 Max Welling提出
- 2014 对比生成模型 (Generative Adversarial Networks/GAN) 由 Ian Goodfellow 和 Yoshua Bengio提出



Geoffrey Hinton,
UofT



Andrew Ng,
Stanford



Max Welling,
UvA



Yoshua Bengio,
Mila



Feifei Li,
Stanford



Ian Goodfellow,
Mila



3 Brasseurs
Crescent
(三位酿酒师)

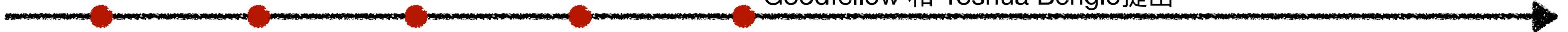
AI简史

2000s - 至今: 深度学习 (Deep Learning)的称霸

- 2006 Geoffrey Hinton 提出 深度置信网络(Depth Belief Networks), 并推广“深度学习 (Deep Learning)”这一概念
- 2008 吴恩达 (Andrew Ng) 团队推动使用GPU进行深度学习计算
- 2012 AlexNet 在 ImageNet图像识别比赛 中获胜 (由 Geoffrey Hinton 团队提出)
- 2013 变分自编码器(Variational Autoencoder/VAE) 由 Diederik Kingman 和 Max Welling提出
- 2014 对比生成模型 (Generative Adversarial Networks/GAN) 由 Ian Goodfellow 和 Yoshua Bengio提出
- 2018 图灵奖授予了三位深度学习先驱
 - Geoffrey Hinton, 多伦多大学
 - 2024年和John Hopfield 因physics for AI, 一起荣获物理诺奖
 - Yoshua Bengio, 蒙特利尔算法研究中心
 - Yann LeCun, 纽约大学 和 Facebook AI 研究院负责人



Yoshua Bengio, Mila



Geoffrey Hinton, UofT



Yann LeCun, NYU



Yann LeCun, NYU



AI简史



Yang Song, Stanford -> OpenAI

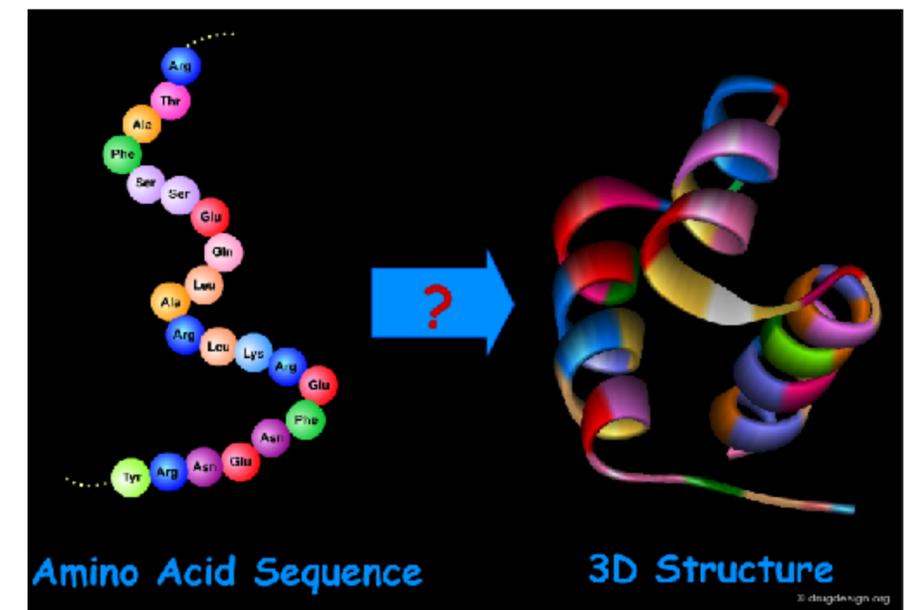
- 2016 AlphaGo (DeepMind) 击败世界围棋冠军
- 2017 Transformer (Google)发布
- 2018 BERT (Google) 与 GPT (OpenAI)发布
- 2019 OpenAI Five (Dota 2 AI) 打败世界冠军OG
- 2019 RoBERTa (Facebook)与 GPT-2 (OpenAI)发布
- 2020 **Noise Conditional Score Network** (Yang Song, Stanford), Denoising Diffusion Probabilistic Models (Google) 与 GPT-3 (OpenAI) 发布
- 2020 AlphaFold2 (DeepMind), AlphaFold 于 2018 年发布, AlphaFold3 于 2024 年发布; AF2的第一作者和最后作者在2024年获化学诺奖。



AlphaGo



OpenAI Five v.s. OG



AlphaFold Series

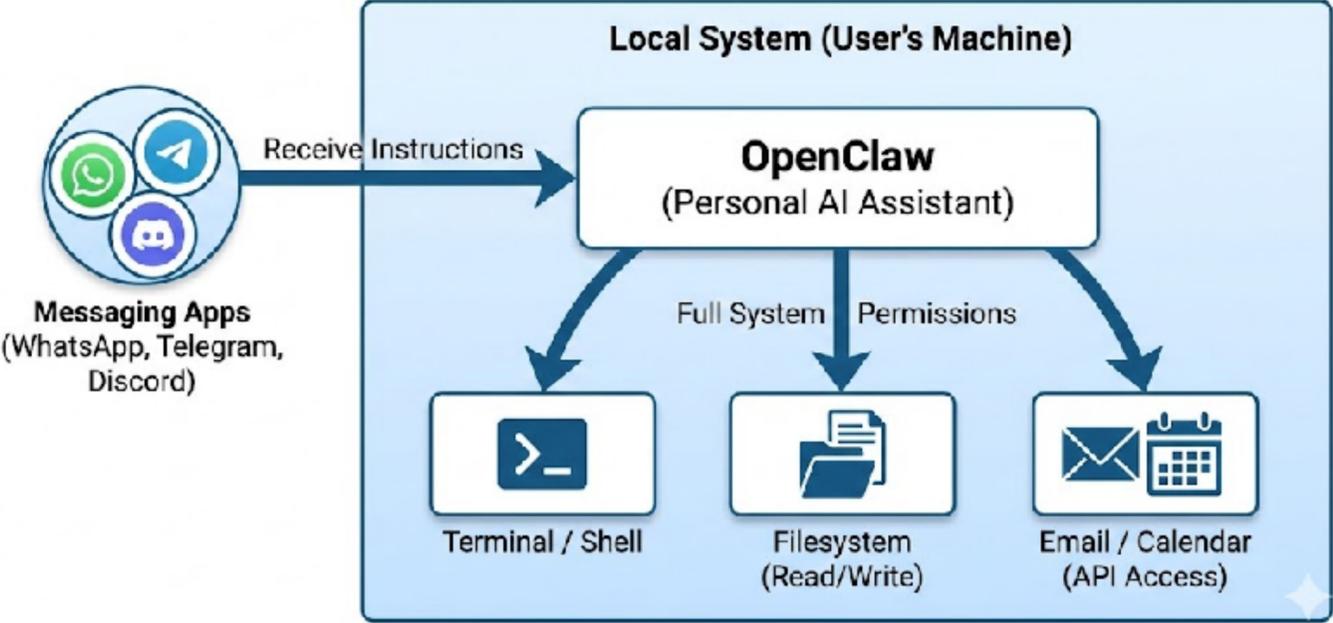
AI简史

- 2022 生成式AI爆发，应用于文本生成图像与问答
 - Imagen (Google), Glide & Dall-E & Dall-E2 (OpenAI), Stable Diffusion (Stability.ai)
 - **ChatGPT** (OpenAI)
- 2023-2024 大语言模型竞赛与多模态模型兴起
 - GPT-4(OpenAI), Claude (Antropic), Gemini (Google), LLaMA (Meta)
 - **DeepSeek** (幻方量化)，工程的极致优化
- 2025-2026.03 更多 AI 产品与应用: AI代理系统
 - Manus? **OpenClaw**?

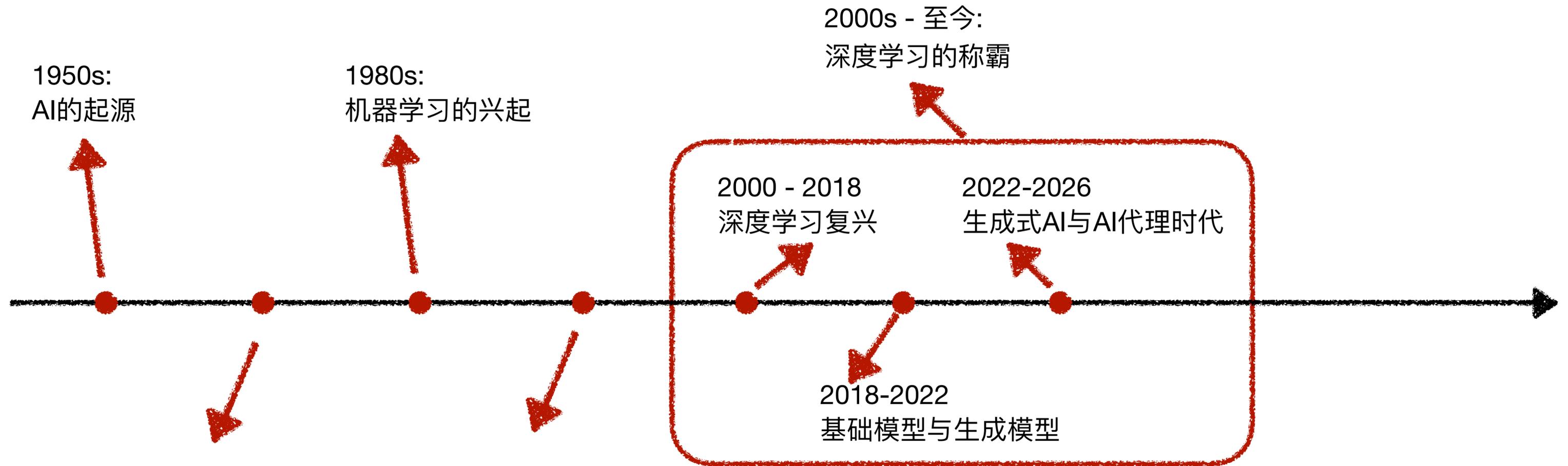


ChatGPT

OpenClaw Local Architecture (Simplified)



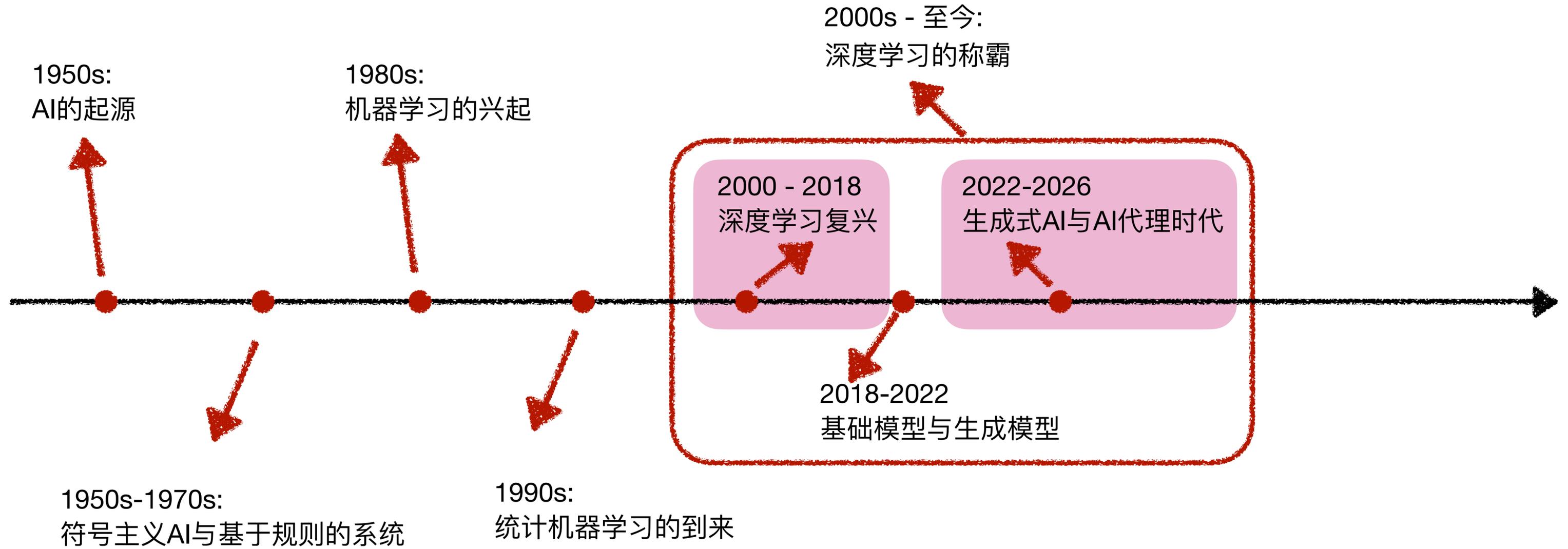
AI简史



Q: 下一步是什么? 先反思一些问题.....

- 我们在对的路子上走么? 大语言模型 (auto-regressive) 和世界模型?
- 现在爆火的方向就是对的方向么?
- 科学和工程如何平衡?
- 如何做出一个成功的产品? 短期和长期?

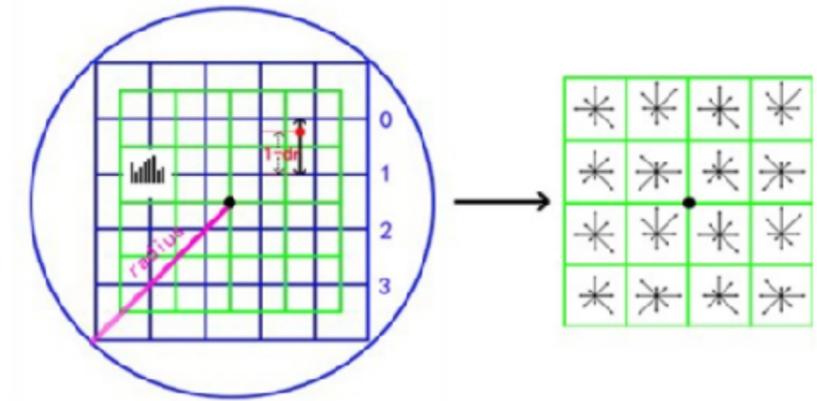
AI简史



1. 深度学习的转角与复兴

机器学习&特征工程 → 深度学习&特征学习

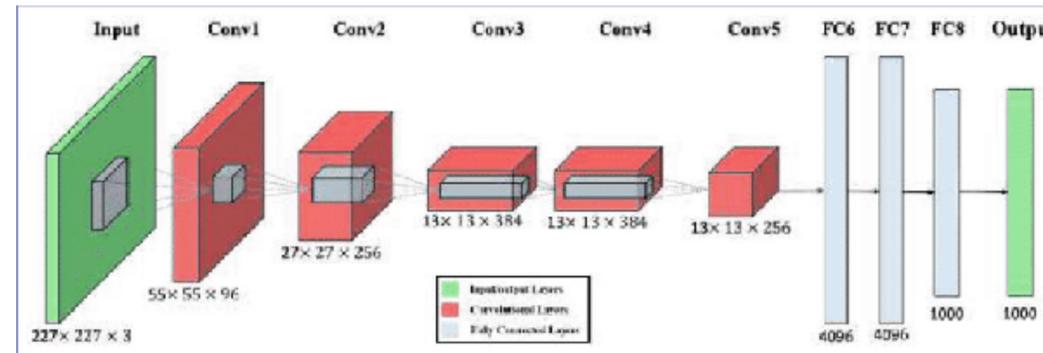
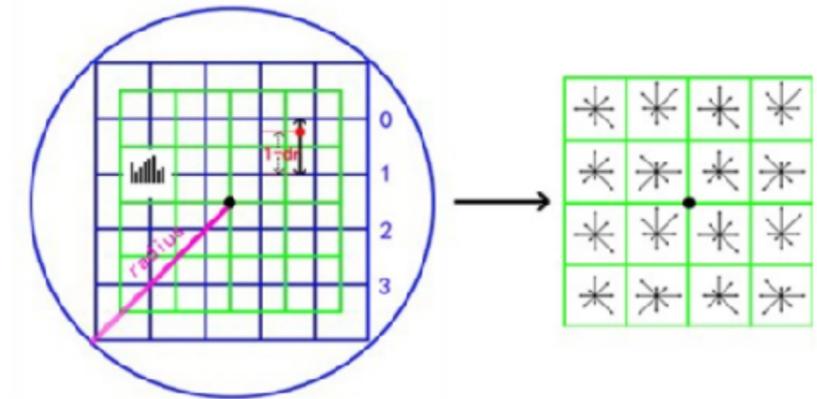
- ImageNet 之前 (<2012)
 - 特征工程: 尺度不变特征变换 (Scale-Invariant Feature Transform/SIFT) 等
 - 模型: 支持向量机 (SVM), 随机森林 (Random Forest), 等



1. 深度学习的转角与复兴

机器学习&特征工程 → 深度学习&特征学习

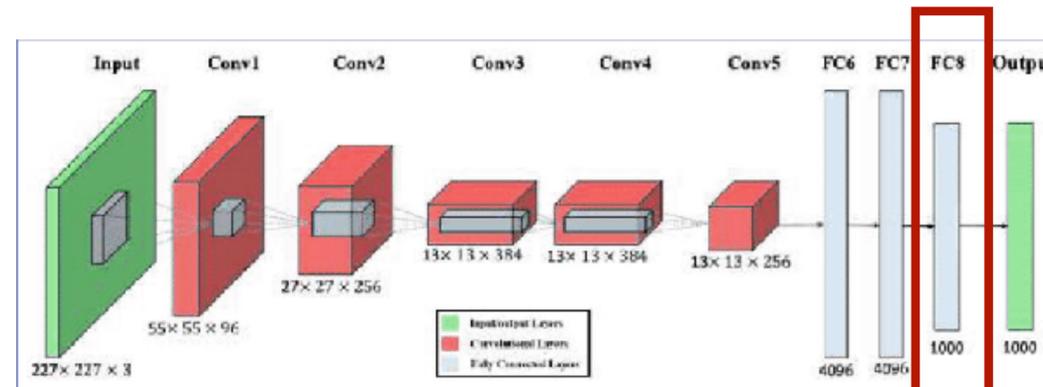
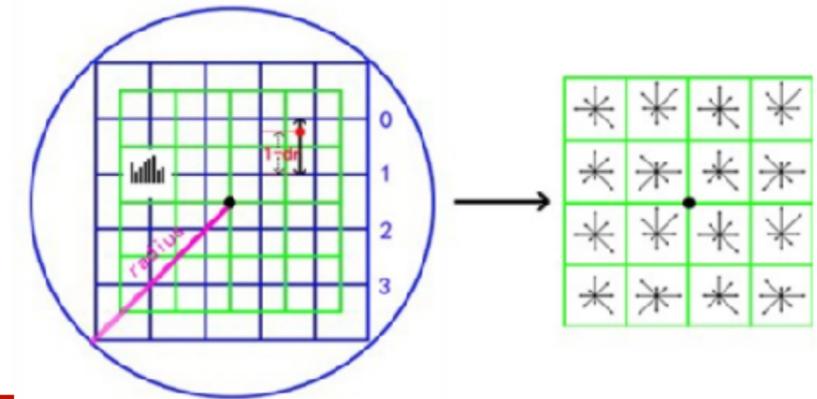
- ImageNet 之前 (<2012)
 - 特征工程: 尺度不变特征变换 (Scale-Invariant Feature Transform/SIFT) 等
 - 模型: 支持向量机 (SVM), 随机森林 (Random Forest), 等
- ImageNet 之后 (≥ 2012)
 - 利用模型进行特征学习: AlexNet



1. 深度学习的转角与复兴

机器学习&特征工程 → 深度学习&特征学习

- ImageNet 之前 (<2012)
 - 特征工程: 尺度不变特征变换 (Scale-Invariant Feature Transform/SIFT) 等
 - 模型: 支持向量机 (SVM), 随机森林 (Random Forest), 等
- ImageNet 之后 (≥ 2012)
 - 利用模型进行特征学习: AlexNet

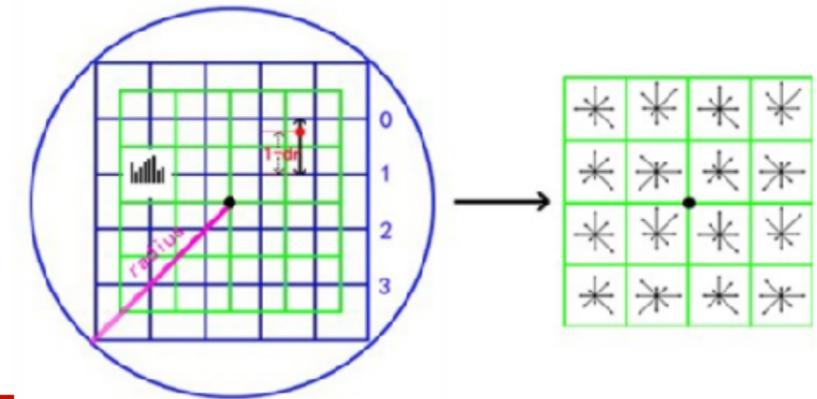


特征 / 表征

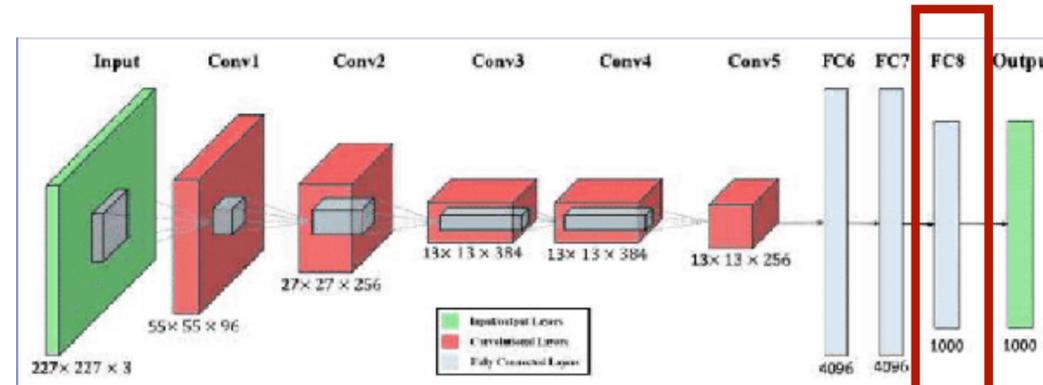
1. 深度学习的转角与复兴

机器学习&特征工程 → 深度学习&特征学习

- ImageNet 之前 (<2012)
 - 特征工程: 尺度不变特征变换 (Scale-Invariant Feature Transform/SIFT) 等
 - 模型: 支持向量机 (SVM), 随机森林 (Random Forest), 等



- ImageNet 之后 (≥ 2012)
 - 利用模型进行特征学习: AlexNet



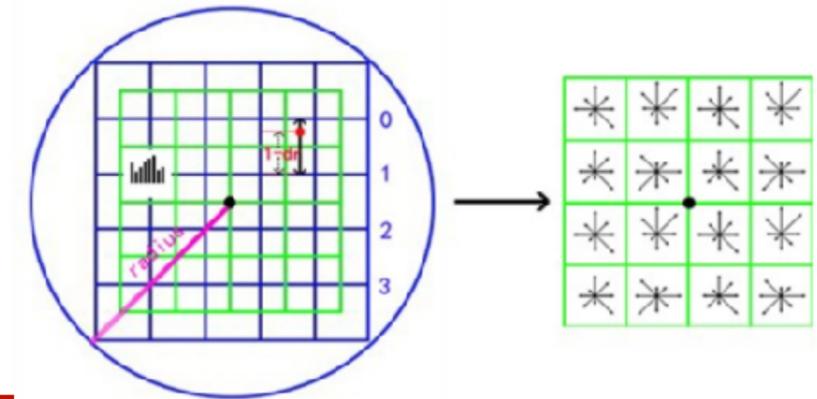
特征 / 表征

- 为什么这个更好?
 - 端到端学习
 - GPU能够提供高性能并行计算
 - 更接近第一性原理: 从数据中进行学习

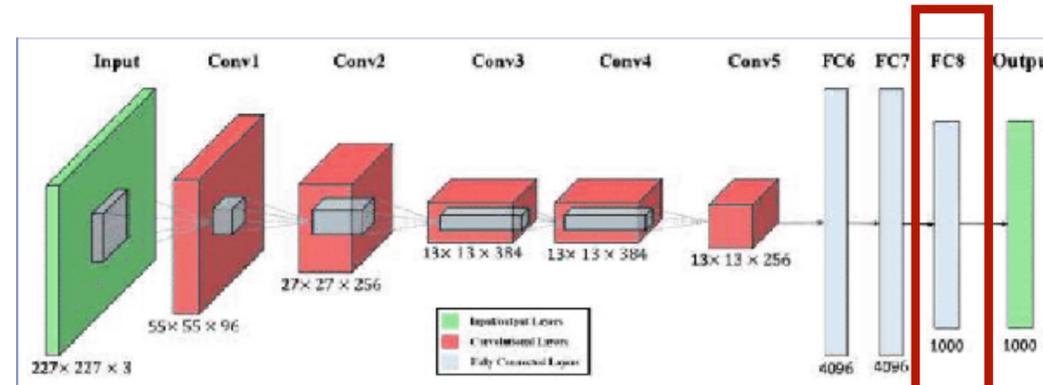
1. 深度学习的转角与复兴

机器学习&特征工程 → 深度学习&特征学习

- ImageNet 之前 (<2012)
 - 特征工程: 尺度不变特征变换 (Scale-Invariant Feature Transform/SIFT) 等
 - 模型: 支持向量机 (SVM), 随机森林 (Random Forest), 等



- ImageNet 之后 (>= 2012)
 - 利用模型进行特征学习: AlexNet



特征 / 表征

- 为什么这个更好?
 - 端到端学习
 - GPU能够提供高性能并行计算
 - 更接近第一性原理: 从数据中进行学习



Yoshua Bengio, Mila

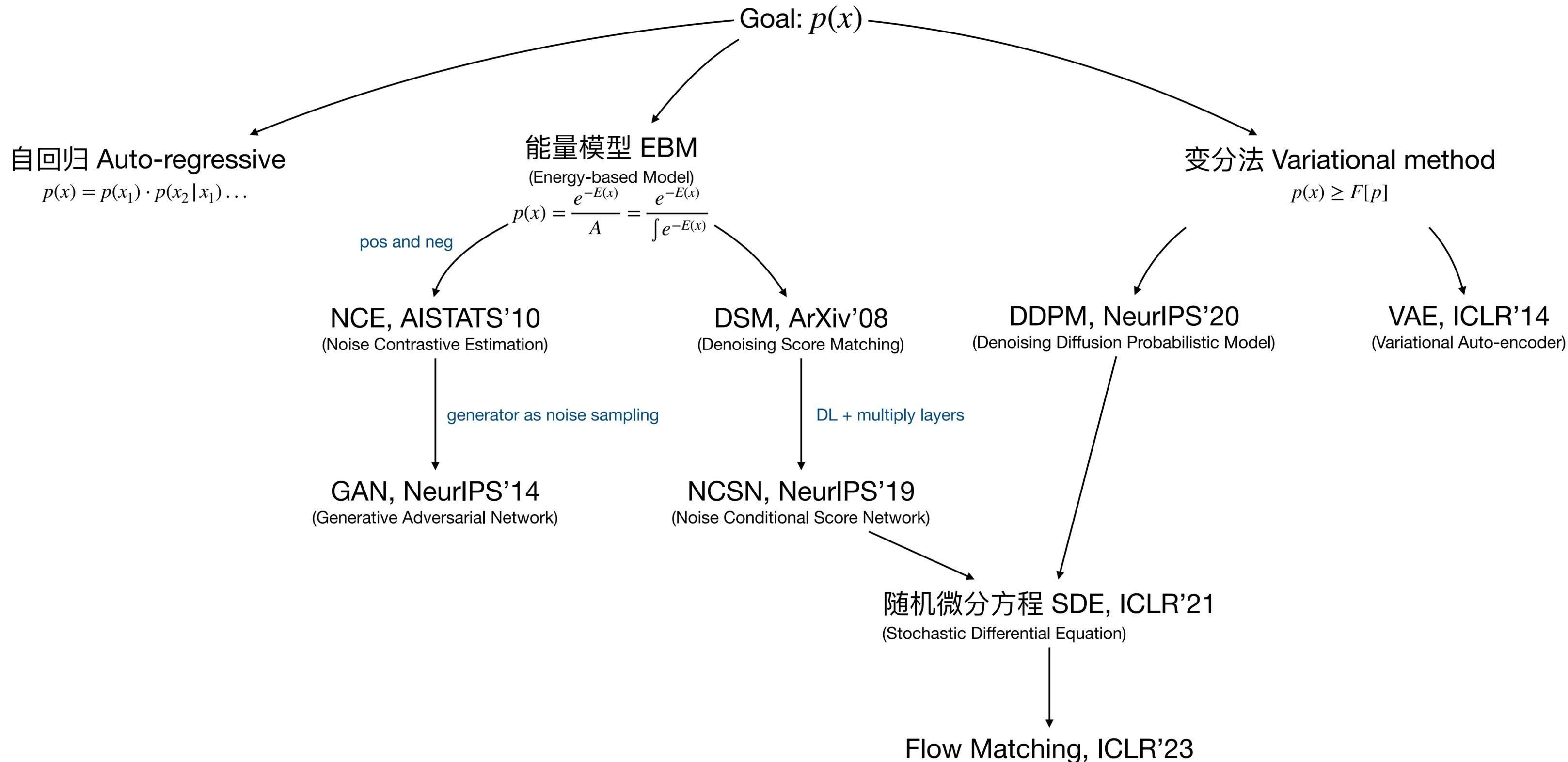


Geoffrey Hinton, UofT



Yann LeCun, NYU

2. 概率估计——除了大语言模型我们还有什么



2. 概率估计——除了大语言模型我们还有什么

Goal: $p(x)$

自回归 Auto-regressive

$$p(x) = p(x_1) \cdot p(x_2|x_1) \dots$$

能量模型 EBM

(Energy-based Model)

$$p(x) = \frac{e^{-E(x)}}{A} = \frac{e^{-E(x)}}{\int e^{-E(x)}}$$

变分法 Variational method

$$p(x) \geq F[p]$$

pos and neg

NCE, AISTATS'10
(Noise Contrastive Estimation)

DSM, ArXiv'08
(Denoising Score Matching)

DDPM, NeurIPS'20
(Denoising Diffusion Probabilistic Model)

VAE, ICLR'14
(Variational Auto-encoder)

generator as noise sampling

DL + multiply layers

大语言模型 LLM
(Large Language Model)

- BERT (**not** AR)
- GPT (AR)
- ...

GAN, NeurIPS'14
(Generative Adversarial Network)
增强式学习 RL
(Reinforcement Learning)

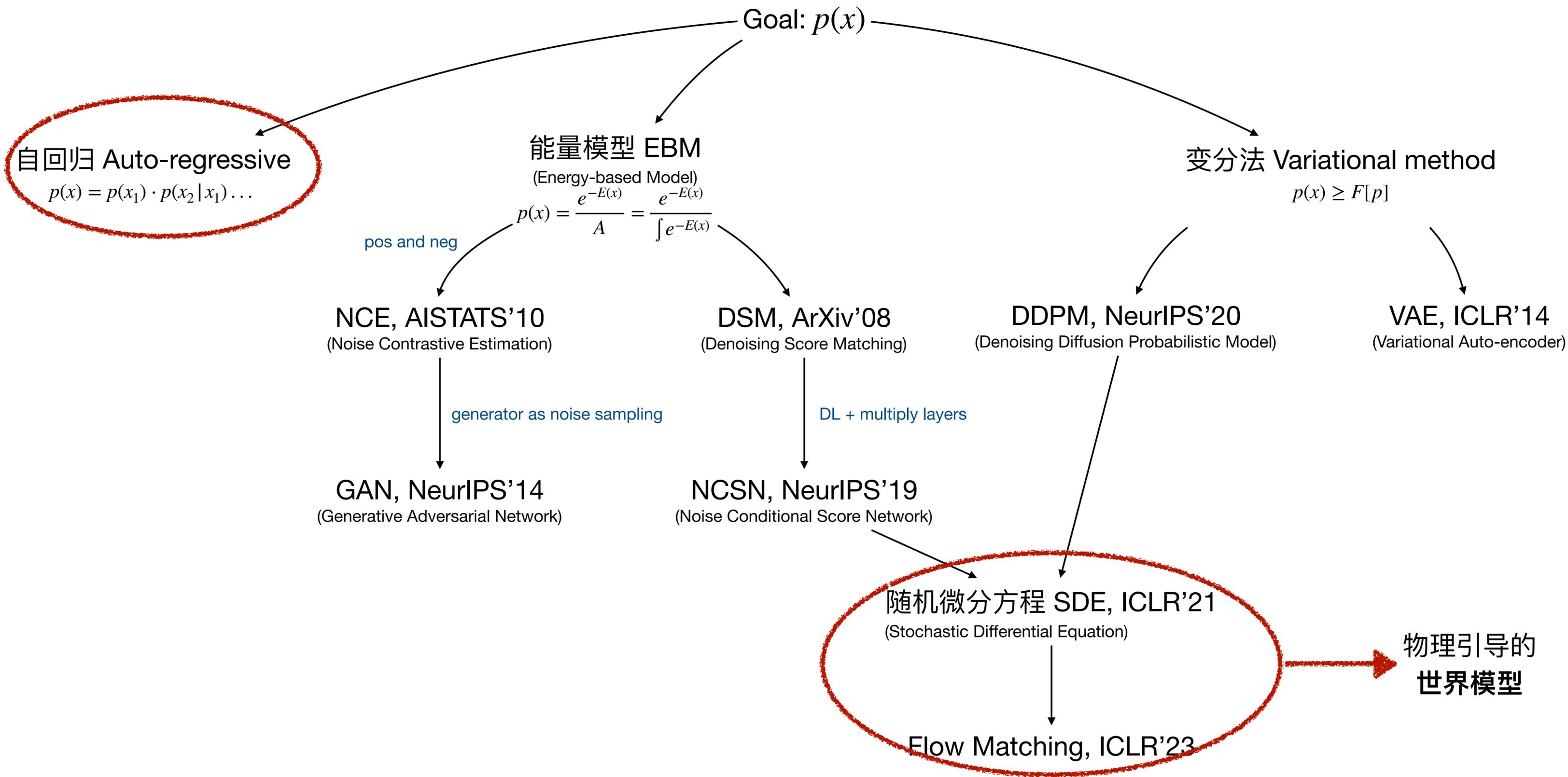


ChatGPT, Llama, DeepSeek,
Qwen 千问

随机微分方程 SDE, ICLR'21
(Stochastic Differential Equation)

Flow Matching, ICLR'23

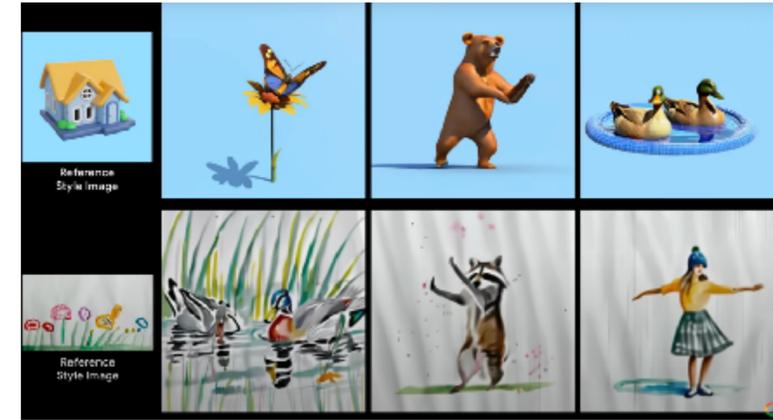
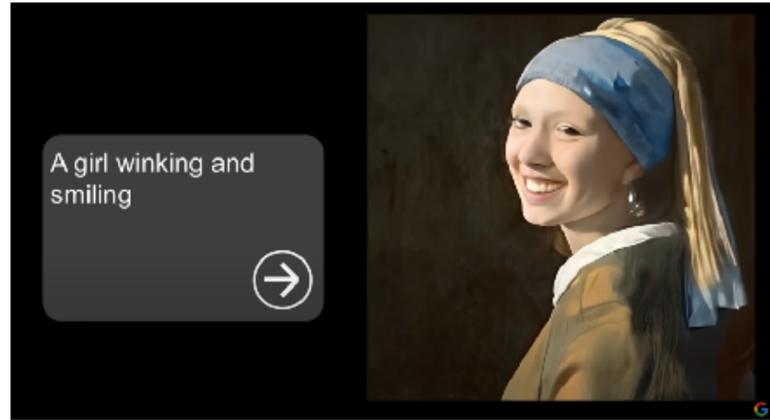
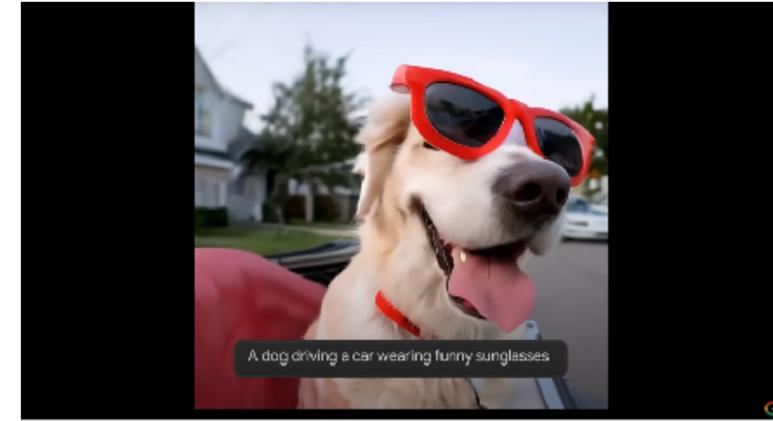
2. 概率估计——除了大语言模型我们还有什么



3. 我们如何利用AI盈利

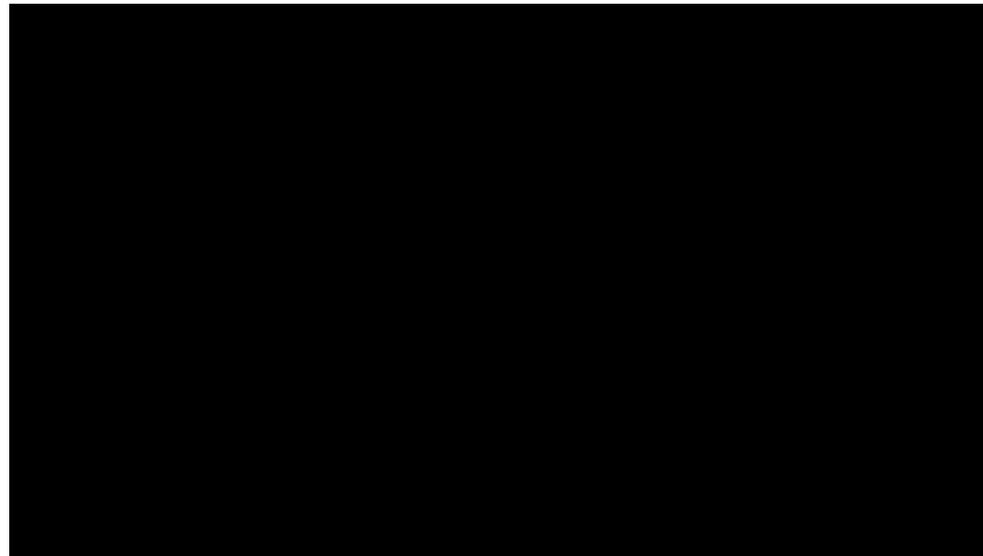
AI for Art 第三产业

Glide, DallE, Imagen, Stable Diffusion, Pika, Sora (OpenAI), Lumiere (Google)



65-year-old Japanese artist, Koya Matsuo (松尾公也)

- Stable Diffusion & Luma Dream Machine
- 用这些软件，通过照片重建视频，来纪念亡妻



3. 我们如何利用AI盈利

AI for Science 第一/二产业

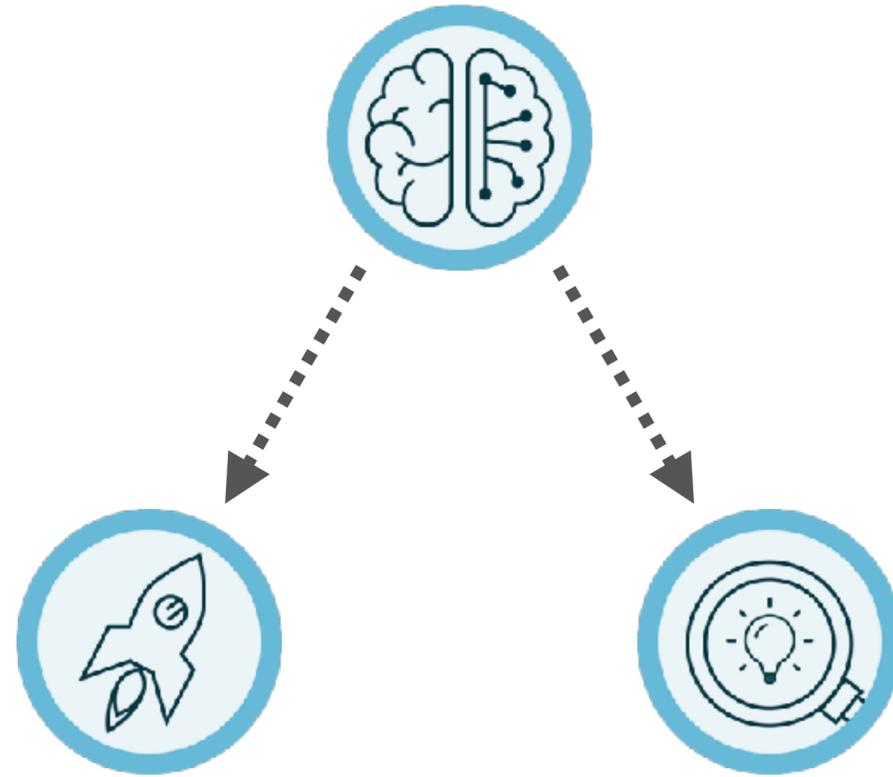
Physics-inspired Pretraining

计算机的起源:

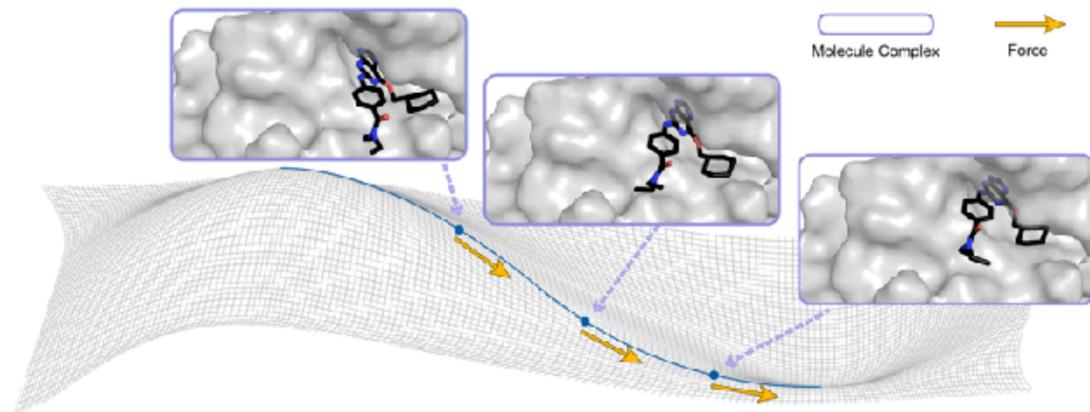
- 破译密码

国家队的进入:

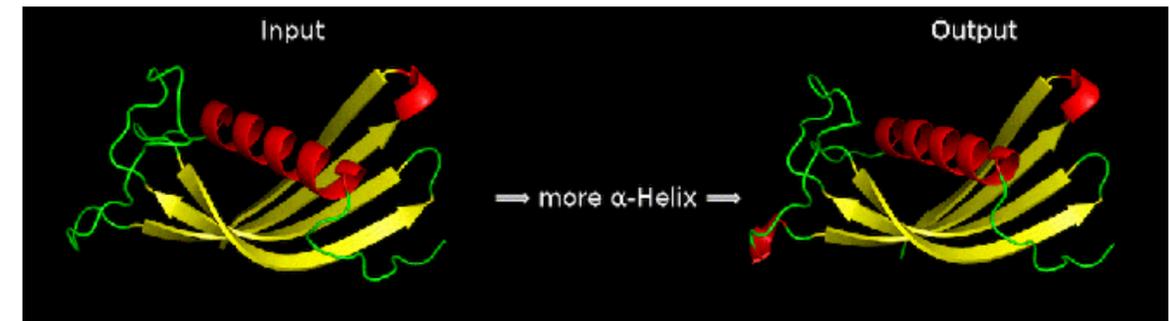
- 具身智能
- AI for science
- 各地的国家实验室



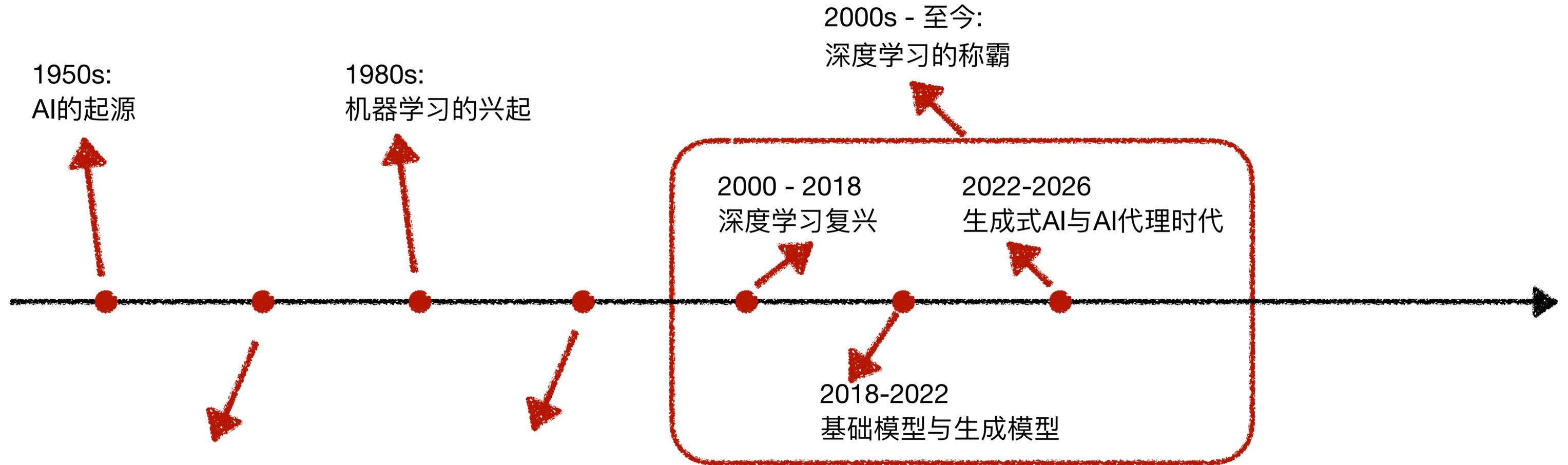
PhysAI for accelerating existing paradigms



GenAI for exploring novel paradigms



AI简史



1950s-1970s:
符号主义AI与基于规则的系统

1990s:
统计机器学习的到来

Q: 下一步是什么? 先反思一些问题.....

- 我们在对的路子上走么? 大语言模型 (auto-regressive) 和世界模型?
- 现在爆火的方向就是对的方向么?
- 科学和工程如何平衡?
 - 如何做出一个成功的产品? 短期和长期?