

历史 AI大事件



01

算法突破



算法突破

- 因为深度学习算法后面的发展越来越快，直接碾压其他机器学习算法，因此现在谈到的 AI 算法，基本上全部都是深度学习算法，而 AI 已经成为了深度学习的代名词了。



算法突破

<https://www.youtube.com/watch?v=TJoWc2l0ohM&pp=ygUQYWxleG5ldCBpbWFnZW5ldA%3D%3D>
<https://www.youtube.com/watch?v=40riCqvRoMs&t=772s&pp=ygUIaW1hZ2VuZXQ%3D>

- 2012年：AlexNet 在李飞飞发起的 ImageNet 挑战赛中夺冠，大幅降低了图像分类错误率，标志着深度学习在图像识别领域的突破。这一事件引发了学术界和工业界对深度学习的广泛关注，促使更多研究和开发资源投入到深度学习领域。

<https://www.youtube.com/watch?v=ZR7tyWUZyHw&pp=ygUSSWFulEdvb2RmZWxsb3cgR0FO>
<https://www.youtube.com/watch?v=7tFBoxex4JE&pp=ygUSSWFulEdvb2RmZWxsb3cgR0FO>

- 2014年：Ian Goodfellow等人提出生成对抗网络（GAN），通过生成器和鉴别器的对抗训练生成高质量数据样本。能够生成逼真的图像、音频等，广泛应用于图像生成、艺术创作等领域。GAN的出现为生成式 AI 建模开辟了新路径，推动了无监督学习的发展，虽然现在生成式 AI 被 Transformer 架构和大模型全面代替了，但是 GAN 才是开启生成式 AI 的鼻祖。



算法突破

<https://www.youtube.com/watch?v=WXuK6gekU1Y&t=3s&pp=ygUOR29vZ2xllEFscGhhR28%3D>

- 2016年：Google 发布的 AlphaGo 击败围棋世界冠军李世石，展示了深度强化学习的强大能力。这一事件证明了 AI 在复杂决策游戏中的强大能力，推动了深度强化学习等技术的发展。

https://www.youtube.com/watch?v=t45S_MwAcOw&pp=ygUSR29vZ2xllFRyYW5zM9ybWVY

- 2017年：Google 提出了 Transformer 架构，通过自注意力机制处理长距离依赖关系，Transformer 成为 NLP 领域的核心技术，为 GPT、BERT 等大语言模型的开发提供了基础。极大地推动了自然语言处理 NLP 的发展，成为现今大模型的基础。Transformer 架构为 NLP 领域带来了革命性变化，推动了机器翻译、文本生成等任务的快速发展。



算法突破

<https://www.youtube.com/watch?v=OFS90-FX6pg&t=950s&pp=ygUNT3BibkFJlEdQVC0xIA%3D%3D>

- 2018年：OpenAI 的 GPT-1 发布， 仅仅使用 Transformer 架构的 Decoder 部分作为网络模型结构，实现了语言生成能力。GPT-1 出现推动了 NLP 在更多行业领域的应用，如智能客服、语言翻译等。为后续 GPT-4， OpenAI 的爆火埋下了很重要的伏笔。
- 2020年：OpenAI 发布 GPT-3， 拥有 1750 亿参数规模，展示了自监督学习在未标记数据训练中的潜力。在文本生成、问答等任务中表现出色，推动了生成式AI的普及。



算法突破

<https://www.youtube.com/watch?v=OFS90-FX6pg&t=950s&pp=ygUNT3BibkFJIEdQVC0xIA%3D%3D>

- 2022年：OpenAI 发布 ChatGPT，展示 ChatGPT 强大的对话和文本生成能力；Stability AI发布Stable Diffusion，推动了文本到图像生成技术的发展。ChatGPT 和 Stable Diffusion 的发布标志着生成式 AI 正式进入主流应用，推动了多模态 AI 的发展。而 GAN 也随着 Stable Diffusion 技术出现而被代替了。

<https://www.youtube.com/watch?v=jay7kise3PI&pp=ygURT3BibkFJIENoYXRHUFQtNG8%3D>

- 2024年：多模态大模型崛起，如OpenAI推出的ChatGPT-4o，能够实时处理和生成文本、音频等多种模态的数据。这一技术的发展为AI在更多领域的应用提供了新的可能性，如智能教育、智能娱乐等。



02

AI 算力与硬件发展



硬件蓬勃发展

- 过去十年，全球科技巨头在 AI 芯片领域的竞争加剧，推动了技术的快速迭代和成本的降低，AI 芯片逐渐成为科技行业的核心竞争力。
- AI 芯片和硬件的发展从专用芯片的崛起到边缘计算的普及，再到量子计算的探索，展现了技术的多样化和应用的广泛性。随着 AI 技术的进一步发展，芯片和硬件继续扮演关键角色，推动 AI 在更多领域的落地和应用。



硬件蓬勃发展

<https://www.youtube.com/watch?v=Ngk8D7f84oE&pp=ygUKR29vZ2xllFRQVQ%3D%3D>

- 2016年：Google 发布了专为机器学习设计的张量处理单元（TPU），用于加速神经网络的推理和训练。TPU 显著提升了 AI 模型的训练效率，尤其是在 AlphaGo 击败李世石的比赛中作为云服务器的执行硬件发挥了关键作用。TPU 的推出标志着 AI 专用芯片的崛起，为后续 AI 芯片的设计提供了重要参考，包括国产的昇腾、寒武纪、壁仞、遂源等。

<https://www.youtube.com/watch?v=fM4JTm9E5os&pp=ygUMTIZJREIBIFZvbHRh>

- 2017年：NVIDIA Volta 架构发布与 AI 芯片元年。NVIDIA 发布了 Volta 架构的 GPU V100，首次引入 Tensor Core，专门优化深度学习任务中的矩阵运算。NVIDIA 也凭借其 GPU 在 AI 芯片市场的领先地位，成为 AI 硬件领域的核心供应商。
- 同时期，因为 2017 年各大科技公司纷纷布局 AI 芯片，因此也被称为 “AI 芯片元年”。



硬件蓬勃发展

https://www.youtube.com/watch?v=9bHzYeH_DjY&pp=ygURYXBwbGUgQTE0IEJpb25pYyA%3D

- 2019年：边缘计算与 AI 芯片的普及。苹果推出 A14 Bionic 芯片，专为边缘设备优化，支持 AI 推理任务。同时，英伟达发布了 Jetson AGX Xavier 开发者套件，推动 AI 在边缘计算中的应用。边缘 AI 芯片的兴起使得 AI 技术能够在智能手机、物联网设备等终端设备上运行，降低了数据传输延迟，提升了隐私保护能力。我们现在手机用到的大部分美颜、瘦脸、指纹解锁、人脸识别开锁等功能，都依赖于芯片内部的 AI 能力。

<https://www.youtube.com/watch?v=ZJXbQPDFVjo&pp=ygUMSGFiYW5hIExhYnMg>

- 2021年：专用AI芯片的多样化。华为发布昇腾 AI 处理器，专注于 AI 推理和训练；谷歌推出 TPU v4，进一步提升AI算力；英特尔 Habana Labs 发布 Gaudi AI 芯片。专用 AI 芯片的多样化满足了不同场景的需求，从数据中心到边缘设备，AI 算力的分布更加均衡，推动了 AI 技术的普及。



硬件蓬勃发展

https://www.youtube.com/watch?v=W7ppd_RY-UE&pp=ygUNZ29vZ2x1IFdpbGxvdmw%3D%3D

- 2023年：量子计算与 AI 结合。谷歌发布量子芯片 Willow，探索量子计算在 AI 领域的应用潜力。量子计算与 AI 的结合为未来 AI 技术的发展提供了新的方向，尤其是在优化算法和解决复杂问题方面展现了巨大潜力，并且提供的高并行计算能力有望解决复杂 AI 问题。

<https://www.youtube.com/watch?v=kZRMshaNrSA&pp=ygUPbnZpZGEgQmxhY2t3ZWxs>

- 2024年：英伟达 Blackwell 架构发布与 AI 芯片的新高度。英伟达发布基于 Blackwell 架构的 B200，拥有 2080 亿个晶体管，推理能力是前代产品的 30 倍，被誉为“史上最强 AI 芯片”，重点支持 FP8 更多低精度格式的计算。Blackwell 架构的推出进一步巩固了英伟达在 AI 芯片市场的领导地位，为生成式 AI 和大模型训练提供了强大的硬件支持。



03

AI 四小龙与六小虎



AI 四小龙

<https://www.youtube.com/watch?v=XaPwbmKZ3Bc&pp=ygUUU2Vuc2VUaW1lIFRIY2hub2xvZ3k%3D>

- AI 四小龙（商汤科技、旷视科技、云从科技、依图科技）是中国 AI 领域的代表性企业，主要聚焦 CV 赛道。它们的兴起得益于深度学习技术的突破、资本的追捧以及安防、金融等行业的巨大需求。然而，四小龙也面临商业化落地难、技术同质化、盈利模式单一等问题，导致估值缩水、上市后股价表现不佳。



AI 四小龙

<https://www.youtube.com/watch?v=XaPwbmKZ3Bc&pp=ygUUU2Vuc2VUaW1lIFRIY2hub2xvZ3k%3D>

- 商汤科技自 2014 年成立以来，经历了从快速崛起、技术领先到面临困境、战略调整的复杂发展历程。其以人脸识别、视频分析等CV 技术起家，凭借强大的技术研发能力迅速成为中国 “AI四小龙” 之首，并在2021年成功在香港交易所上市。
- 然而，近年来由于业务增长停滞、传统AI业务收入下滑以及持续亏损等问题，商汤科技市值大幅缩水，较上市时蒸发超过八成，2024年上半年亏损净额为24.77亿元。



AI 四小龙

<https://www.youtube.com/watch?v=0zqvw34AHV&t=81s&pp=ygURTWVndmlpIFRIY2hub2xvZ3k%3D>

- 旷视自 2011 年成立以来，同样经历了从快速崛起、面临困境到战略转型的波折。早期凭借 Face++ 平台在人脸识别领域取得先发优势，成功切入金融、安防等行业，2016 年达到巅峰。然而，2019 年公司上市受阻，AI 业务也面临挑战。最后撤回科创板上市申请，相关技术人员也转到大模型 6 小虎里面。



大模型 6 小虎

- 2023 年是大模型元年，也是在 2023 年，国内形成了大模型“六小虎”的格局，主要包括百川智能、零一万物、智谱AI、MiniMax、月之暗面与阶跃星辰，他们的发展现状呈现出多方面的特点：
- **技术瓶颈：**大模型技术发展从 2024 年进入深水区，迭代速度变慢，多模态仍处于攻坚早期。OpenAI 在 GPT-4 之后的技术报告不再公开，很多基础大模型厂商失去参考方向，现在主要靠开源牵引整体技术路线发展。
- **市场困境：**2024 年下半年，六小虎的口碑急转直下，潜在买家减少，除了国资外，难得有其他潜在买家或者下一轮资本介入。在业务拓展方面，国内 C 端用户忠诚度不高，B 端市场盈利困难。



大模型 6 小虎

- **市场竞争：**大厂如字节等开始强势崛起，利用流量、算法等优势，在To B和To C领域对六小虎形成碾压。
- **未来展望：**2025 年以后，ZOMI 预计基础模型企业进入赛道的难度将增大，新入局者很难拿到融资。未来的发展将主要集中在基础模型与行业的结合，寻找商业化落地的途径，因此大家预测 2025 年是 AI Agent 的元年。
- 但是 AI Agent 是什么？ AI Agent 怎么定义？ AI Agent 的技术路线跟 RAG 、 RL、大模型之间的怎么结合？



总结



XXX

- 过去 10 年，AI 算法从深度学习（如 AlexNet、ResNet）到 Transformer 架构（如 BERT、GPT）的突破，推动了自然语言处理、计算机视觉等领域的革命性进展，生成式 AI（如 ChatGPT、Stable Diffusion）成为新焦点。
- AI 专用芯片（如 GPU、TPU、NPU）快速发展，英伟达、谷歌、华为等公司推动算力大幅提升，支持更大规模模型训练和实时推理，边缘计算和低功耗芯片也逐渐普及。
- AI 初创企业（如 OpenAI、DeepMind、幻方量化、商汤科技）在资本和技术红利下迅速崛起，推动 AI 技术商业化落地，传统 AI 初创面临盈利压力和市场整合，行业进入理性发展阶段。而大模型初创则进入新一轮的竞争。





Thank you

把AI系统带入每个开发者、每个家庭、
每个组织，构建万物互联的智能世界

Bring AI System to every person, home and
organization for a fully connected,
intelligent world.

Copyright © 2024 XXX Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. XXX may change the information at any time without notice.



GitHub <https://github.com/chenzomi12/AIFoundation>