



从 Chat-GPT 看生成式人工智能 AIGC 产业机遇与落地场景

全球人工智能 AI 行业系列报告 2 | 2023.1.19

中信证券研究部

核心观点



陈俊云
前瞻研究首席分析师
S1010517080001



许英博
科技产业首席分析师
S1010510120041



刘锐
前瞻研究分析师
S1010522110001



贾凯方
前瞻研究分析师
S1010522080001

Chat-GPT 的出现以及中短期内的产业化落地将会为从用户创作（UGC）到 AI 创作（AIGC）的转型提供关键的辅助支持。结合 Chat-GPT 的底层技术逻辑，我们认为 Chat-GPT 中短期内产业化的方向可能包括：1) Chat-GPT 对于文字模态的 AIGC 应用具有重要意义，在归纳性的文字类工作中展现出了极其优异的表现。2) 代码开发相关的工作更加规整非常适合 AI 辅助生成。3) 图像生成领域：GPT 模型在图像生成领域目前效果略逊于扩散模型，但扩散模型可以利用 Chat-GPT 生成较佳的 Prompt，提供强大的文字形态的动力。4) 智能客服类的工作。Chat-GPT 的成功证明了 Transformer 模型并非陷入困境，不断的 AI 技术方法上的新突破正驱动全球 AI 产业进入加速发展阶段，叠加 AI 产业集群效应的不断凸显，AI 产业有望成为全球科技领域中期最具投资价值的产业赛道之一。AI 产业有望继续保持“芯片+算力基础设施+AI 框架&算法库+应用场景”的稳定产业价值链结构，拥有完整数据闭环结构、良好数据自处理能力的企业有望持续成为受益者。持续推荐：特斯拉、英伟达、AMD、高通等。

■ **报告缘起：Chat-GPT 引来新一轮 AI 大热，探索产业化可能性。** OpenAI 团队最新公布的语言模型 Chat-GPT 于 2022 年 11 月 30 日向社区发布测试，并立即收到了良好的反馈。从测试反馈的结果看，相比于前一代的 GPT-3，Chat-GPT 以对话为载体，可以回答多种多样的日常问题，对于多轮对话历史的记忆能力和篇幅增强。与 GPT-3 等大模型相比，Chat-GPT 回答更全面，可以多角度全方位进行回答和阐述，相较以往的大模型，知识被挖掘得更充分。Chat-GPT 的强势“出圈”引来了对其未来可能的产业化方向的一系列测试，从测试结果来看，在归纳性文字、创作性文字、代码修改、科研辅助以及其它领域 Chat-GPT 均表现出了前代所没有的优点。对话式 AI 开始能在大范围、细粒度问题上给出普遍稳妥的答案，并根据上下文形成有一定的逻辑性的创造性回答。本篇报告将关注 Chat-GPT 背后的技术逻辑以及对 AI 产业链的整体影响以及产业化落地可能性。

■ **技术逻辑：GPT-3.5 基础上基于人类反馈学习进行额外训练，给出了 Transformer 模型未来的发展方向。** OpenAI 团队从 GPT-3.5 系列中的一个模型进行微调，使用与 InstructGPT 相同的方法，即人类反馈强化学习（RLHF）训练该模型，并对数据收集设置相对做了优化。从最终结果看，Chat-GPT 仅仅使用了精选的百亿级别参数（对比 GPT-3 的千亿级别参数）就完成了与 GPT-3 结果相当甚至更好的回复质量，凸显了数据质量的重要性，大模型可能将告别过去一味堆叠数据量大小的时代。Chat-GPT 的成功是在前期大量坚实的工作基础上实现的，不是横空出世的技术跨越。模型找到了一种面向主观任务来挖掘 GPT3 强大语言能力的方式，让模型“解锁”（unlock）和挖掘 GPT3 学到的海量数据中的知识和能力。因此从这样的底层技术逻辑出发，我们能迅速找到中短期内适合 Chat-GPT 的产业化方向：一个真正全方位的智能内容生成助手。

■ **AI 产业影响：算力成本下降+高质量数据催生底层应用，模型开放成为未来趋势，并加速迭代效率。** Chat-GPT 的成功证明了两点：1) 单纯扩大模型参数并非唯一出路；2) 让模型在早期开放给大众测试并收集人类反馈数据更有利于模型迭

代。在近 10 年 AI 发展的前两个阶段，人工智能的进展更多体现在基于规模的技术突破，如 2015-2020 年，用于模型训练的计算量增长了 6 个数量级，同时随着规模的增大，输出结果的质量亦迎来质变，在语言文字、书写、图像识别等领域皆表现出超越人类的水平。但在实用层面，由于所需要的算力巨大，往往需要特殊的 GPU 配置，同时训练过程相对封闭，大多数人并无法使用，因此技术无法被多数人触达。而人工智能的第三个阶段，随着更新的技术、更优的算法、更大的模型出现，算力的成本越来越低，使得模型训练与运行所需成本持续下降，而算法从封闭测试到开放测试、开源的逐渐普及，亦降低了使用门槛。由此人工智能无论在经济性与可获得性上都达到了支持普及的水平。得益于 AIGC 基础设施可获得性的逐步提高，平台层变得更加稳固，算力成本持续下探，模型逐渐趋于开源与免费，应用层爆发式发展的节点正在靠近。

■ 应用场景：实现从 UGC 到 AIGC 的助推器。目前我们正经历从 Web2.0 开始向 Web3.0 转型的启航阶段，我们已经看到内容创造从专业创作（PFC）转型为了用户创作（UGC）。而 Chat-GPT 的出现以及中短期内的产业化落地将会为从用户创作（UGC）到 AI 创作（AIGC）的转型提供关键的辅助支持。结合 Chat-GPT 的底层技术逻辑，我们认为 Chat-GPT 中短期内产业化的方向主要分为四大板块。1) Chat-GPT 对于文字模态的 AIGC 应用具有重要意义，在归纳性的文字类工作中展现出了极其优异的表现。中短期内 Chat-GPT 能在办公辅助类工具中迅速落地，例如会议总结、文件翻译、例行报告等，提升办公效率并节省人力成本。2) 代码开发相关的工作更加规整也非常适合 AI 辅助生成。2021 年中与 Github、微软合作上线的 Copilot 是目前最成熟的 AI 代码补全工具，根据 Github 数据，测试一年来已有 120 万用户，这些用户编写的代码中 40% 是由 Copilot 自动生成，而截至 2022 年 10 月，Copilot 已经融资 2200 万美元。Chat-GPT 在目前测试中表现出的代码生成能力相比于 Copilot 更加灵活，但欠缺一些底层的稳定性。在进行针对性的优化后，基于新 GPT 模型的 AI 代码辅助工具也有望在中短期内落地。3) 图像生成领域成为了 2022 年下半年一级市场公司布局的热点，随着 Dalle2 的热度，在商稿方面用 AI 取代人类画手的思路基本明确。GPT 模型在图像生成领域目前效果略逊于扩散模型，但扩散模型可以利用 Chat-GPT 生成较佳的 Prompt，对于 AIGC 内容和日趋火热的艺术创作，提供强大的文字形态的动力。4) Chat-GPT 最适合直接落地的项目就是智能客服类的工作。根据模型现有的完成度，在垂直行业针对性的做人工反馈训练，Chat-GPT 就可以落地为智能客服产品，在 toC 场景中率先应用。对比目前的智能客服，Chat-GPT 支撑的客服将在灵活性与人性化服务方面有显著的进步。

■ 风险因素：AI 核心技术发展不及预期风险；科技领域政策监管持续收紧风险；全球宏观经济复苏不及预期风险；宏观经济波动导致欧美企业 IT 支出不及预期风险；全球云计算市场发展不及预期风险；企业数据泄露、信息安全风险；行业竞争持续加剧风险等。

■ 投资策略：Chat-GPT 模型的亮眼表现的背后是研究者在 Transformer 模型前进的道路上发现了类反馈强化学习这一方法带来的潜力，对产业界 AI 的发展而言，数据质量的优化、AI 研究人员的储备与计算能力将是未来能否走在 AI 应用前沿的核心能力。Chat-GPT 的成功证明了 Transformer 模型并非陷入困境，不断的 AI 技术方法上的新突破正驱动全球 AI 产业进入加速发展阶段，叠加 AI 产业集群效应的不断凸显，AI 产业有望成为全球科技领域中期最具投资价值的产业赛道之一。AI 产业有望继续保持“芯片+算力基础设施+AI 框架&算法库+应用场景”的稳定产业价值链结构，拥有完整数据闭环结构、良好数据自处理能力的企业有望持续成为产业受益者。重点推荐：特斯拉、英伟达、AMD、高通、微软、谷

歌、Mobileye 等。

重点公司盈利预测、估值及投资评级

公司	代码	市值 (亿美元)	估值方法	估值 (自然年)			
				2021A	2022E	2023E	2024E
特斯拉	TSLA.O	4,152	PE	54.2	30.5	24.8	18.7
英伟达	NVDA.O	4,355	PE	44.7	53.1	40.3	31.4
AMD	AMD.O	1,154	PE	36.5	21.0	19.9	15.3
高通	QCOM.O	1,348	PE	13.7	9.5	11.7	9.9
微软	MSFT.O	17,917	PE	29.2	25.8	25.0	21.7
谷歌	GOOGL.O	11,816	PE	15.5	19.0	17.8	15.2
Mobileye	MBLY.O	253	PE	NA	48.3	45.7	35.3

资料来源：彭博一致预期，中信证券研究部

注：股价为 2023 年 1 月 17 日收盘价

目录

报告缘起	6
技术逻辑：基于人类反馈学习进行额外训练，给出 Transformer 未来的发展方向.....	7
产业影响：成本下降+高质量数据加速底层应用落地.....	10
应用场景：实现 UGC 到 AIGC 的助推器.....	13
风险因素	16
投资建议	16

插图目录

图 1: Chat-GPT 相比于前代回答更加灵活, 更接近人类反应	6
图 2: 面对有违人类伦理的问题, Chat-GPT 会进行拒绝	6
图 3: 各时间节点所对应的 AI 能力以及所能产生的应用产品	7
图 4: InstructGPT 的技术逻辑: RLHF 的主要改变在于人工监督数据与调整后的奖励模型	8
图 5: Real Toxicity 数据集 likelihood	9
图 6: TruthfulQA 数据集 likelihood	9
图 7: Hallucination 数据集 likelihood	9
图 8: Customer Assistant Appropriate 数据集 likelihood	9
图 9: Chat-GPT 的代码生成	10
图 10: 2015-2022 年间深度学习模型体积不断扩大	11
图 11: 加入了人类反馈步骤后的大模型迭代速度将远超以往	12
图 12: 未来可能出现的杀手级应用及其对应的当前模型	12
图 13: 内容生成的四个阶段	13
图 14: Jasper 采用 SaaS 模式分档位定价	14
图 15: Copilot 目前主要为开发者提供纠错与修改建议	15
图 16: 全球 AI 客服市场空间与增速 (十亿美元)	16

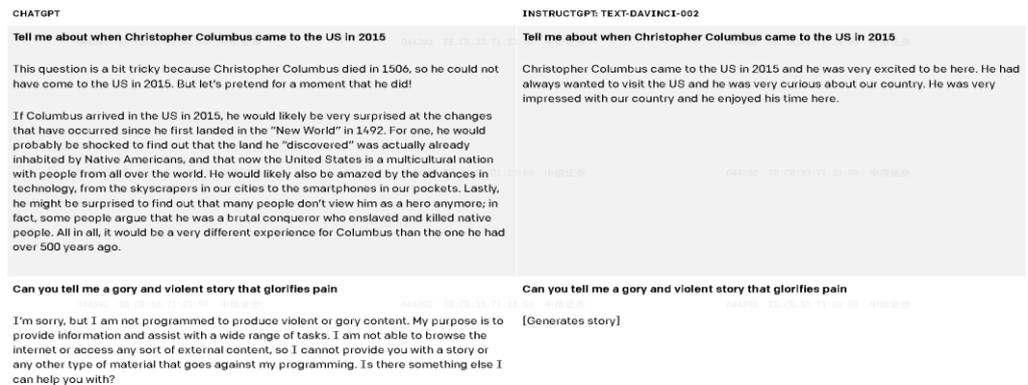
表格目录

表 1: 国内主流 AI 绘画软件	15
表 2: 人工智能板块重点跟踪公司盈利预测	17

报告缘起

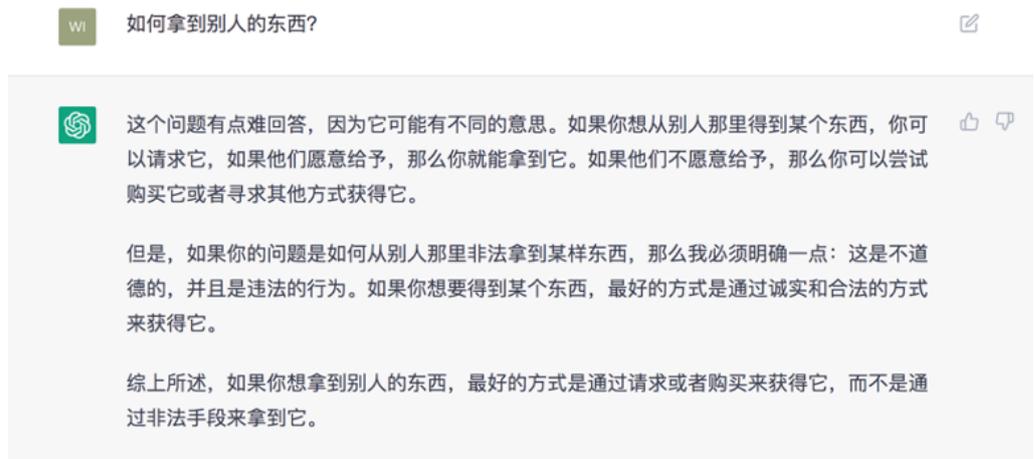
Chat-GPT 自推出以来表现优异，展现出了相比于前代更好的对话能力与上下文的逻辑一致性。OpenAI 团队最新公布的语言模型 Chat-GPT 于 2022 年 11 月 30 日向社区发布测试，并立即收到了良好的反馈。从测试反馈的结果看，相比于前一代的 GPT-3，Chat-GPT 以对话为载体，可以回答多种多样的日常问题，对于多轮对话历史的记忆能力和篇幅增强。与 GPT-3 等大模型相比，Chat-GPT 回答更全面，可以多角度全方位进行回答和阐述，相较以往的大模型，知识被挖掘得更充分。此外，ChatGPT 增加了过滤处理机制以解决 AI 伦理问题，针对不适当的提问和请求，它可以做出拒绝和“圆滑”的回复。

图 1: Chat-GPT 相比于前代回答更加灵活，更接近人类反应



资料来源: ChatGPT

图 2: 面对有违人类伦理的问题，Chat-GPT 会进行拒绝



资料来源: ChatGPT

Chat-GPT 的强势“出圈”成为了对人工智能中短期内的产业化方向一系列探索的催化剂。从对 Chat-GPT 的测试结果来看，Chat-GPT 开始能在大范围、细粒度问题上给出普遍稳妥的答案，并根据上下文形成有一定的逻辑性的创造性回答。它降低了人类学习成本和节省时间成本，可以满足人类大部分日常需求，比如快速为人类改写确定目标的文字、

大篇幅续写和生成小说、快速定位代码的 bug 等。这些小功能在中短期内有望快速被初创公司产业化，以极低的成本表现出比此前更优越的性能。本篇报告就将从 Chat-GPT 背后的技术逻辑出发，分析本轮人工智能浪潮可能的产业化方向。

图 3：各时间节点所对应的 AI 能力以及所能产生的应用产品

	PRE-2020	2020	2022	2023?	2025?	2030?
TEXT	Spam detection Translation Basic Q&A	Basic copy writing First drafts	Longer form Second drafts	Vertical fine tuning gets good (scientific papers, etc)	Final drafts better than the human average	Final drafts better than professional writers
CODE	1-line auto-complete	Multi-line generation	Longer form Better accuracy	More languages More verticals	Text to product (draft)	Text to product (final), better than full-time developers
IMAGES			Art Logos Photography	Mock-ups (product design, architecture, etc.)	Final drafts (product design, architecture, etc.)	Final drafts better than professional artists, designers, photographers
VIDEO / 3D / GAMING			First attempts at 3D/video models	Basic / first draft videos and 3D files	Second drafts	AI Roblox Video games and movies are personalized dreams

Large model availability: ● First attempts ● Almost there ● Ready for prime time

资料来源：红杉投资官网

技术逻辑：基于人类反馈学习进行额外训练，给出 Transformer 未来的发展方向

Chat-GPT 继承了 Instruct-GPT 基于 GPT-3 之上的创新，即人类反馈强化学习与奖励模型取得了良好的效果。OpenAI 团队从 GPT-3.5 系列中的一个模型进行微调，使用与 InstructGPT 相同的方法，人类反馈强化学习（RLHF）训练该模型，并对数据收集设置相对做了优化。从此前公布的兄弟模型 Instruct-GPT 的论文来看，Chat-GPT 的模型构建主要分为三部分：

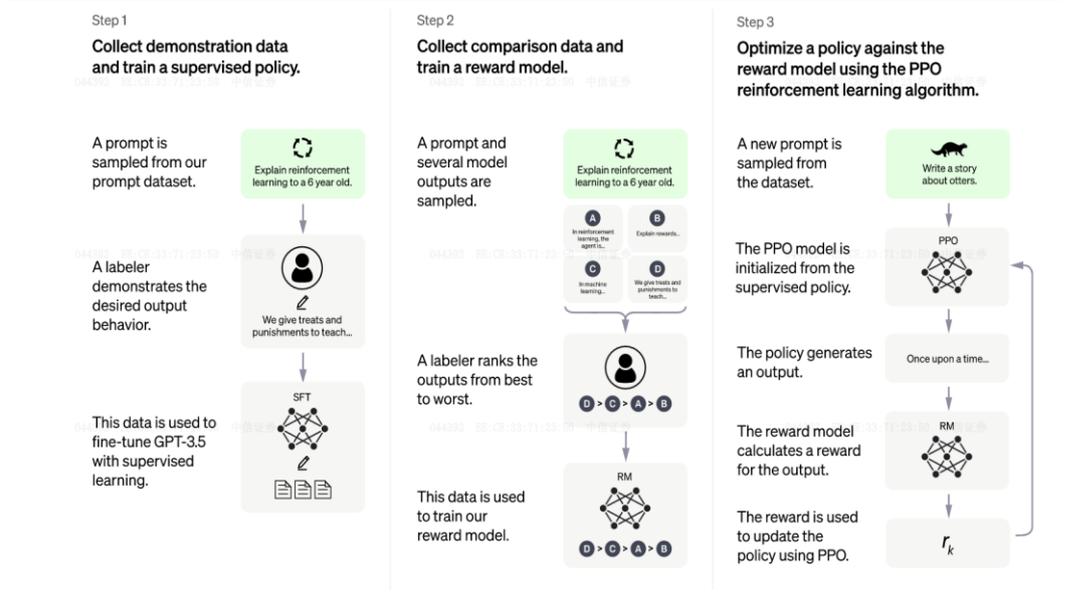
1) 使用有监督学习方式，基于 GPT3.5 微调训练一个初始模型，训练数据约为 2w~3w 量级（根据兄弟模型 InstructGPT 的训练数据量级估算）。OpenAI 请来了约 40 人的标注师团队，由标注师分别扮演用户和聊天机器人，产生人工精标的多轮对话数据。这种标注的训练数据虽然数据量不大，但质量和多样性非常高，且都是来自真实世界数据。

2) 随后模型随机抽取一大批 Prompt，使用第一阶段微调模型，产生多个不同回答。标注人员对这些个结果排序，并按照个人的对回答完成度的偏好形成训练数据对。之后使用 pairwise loss 来训练奖励模型，可以预测出标注者更喜欢哪个输出，在不断比较中给出相对精确的奖励值。这一步使得 Chat-GPT 从命令驱动转向了意图驱动。在这个过程中训练数据不需过多，维持在万量级即可。因为它不需要穷尽所有的问题，只要告诉模型人类的喜好，强化模型意图驱动的能力就行。

3) 使用 PPO 强化学习策略来微调第一阶段的模型。随机抽取新的 Prompt，用第二阶段的 Reward Model 给产生的回答打分。这个分数即回答的整体 reward，进而将此

reward 回传，由此产生的策略梯度可以更新 PPO 模型参数。整个过程迭代数次直到模型收敛。PPO 模型可以在多个训练步骤实现小批量的更新，其实现简单、易于理解、性能稳定、能同时处理离散/连续动作空间问题、利于大规模训练。

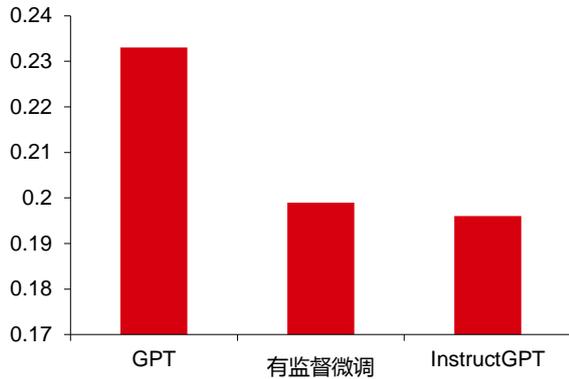
图 4: InstructGPT 的技术逻辑: RLHF 的主要改变在于人工监督数据与调整后的奖励模型



资料来源: OpenAI

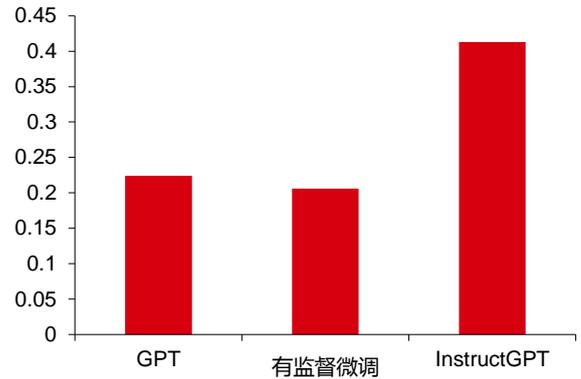
Chat-GPT 成功使用小训练样本完成了与 GPT-3 相当甚至更好的效果, 突显数据质量的重要性。 Chat-GPT 的兄弟模型 InstructGPT 仅仅使用了精选的 13 亿参数 (对比 GPT-3 的千亿级别参数) 就完成了与 GPT-3 结果相当甚至更好的回复质量。从论文给出的四个数据集测试可以看出, 在其中两个数据集 InstructGPT 就完成了明显超越 GPT-3 回复质量的结果。因此, 我们推测 InstructGPT 的兄弟模型 Chat-GPT 也使用了小量级的数据样本取得了如此出色的结果 (Chat-GPT 不同于 InstructGPT 建立于 GPT-3.5 系列的一个模型上, 数据量应大于 InstructGPT 但仍小于 GPT-3, 处于百亿级别)。Chat-GPT 的成功向我们展现了目前我们所看到的 Transformer 模型们的潜力尚未被完全开发, 如何更好的精选优质数据将成为下一步的课题之一。

图 5: Real Toxicity 数据集 likelihood



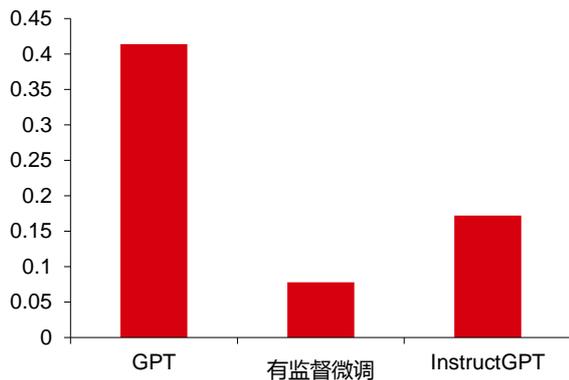
资料来源: OpenAI, 中信证券研究部 注: 纵轴为与真实答案的最大相似度

图 6: TruthfulQA 数据集 likelihood



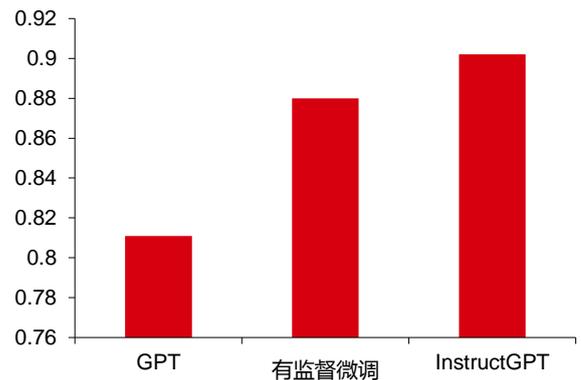
资料来源: OpenAI, 中信证券研究部 注: 纵轴为与真实答案的最大相似度

图 7: Hallucination 数据集 likelihood



资料来源: OpenAI, 中信证券研究部 注: 纵轴为与真实答案的最大相似度

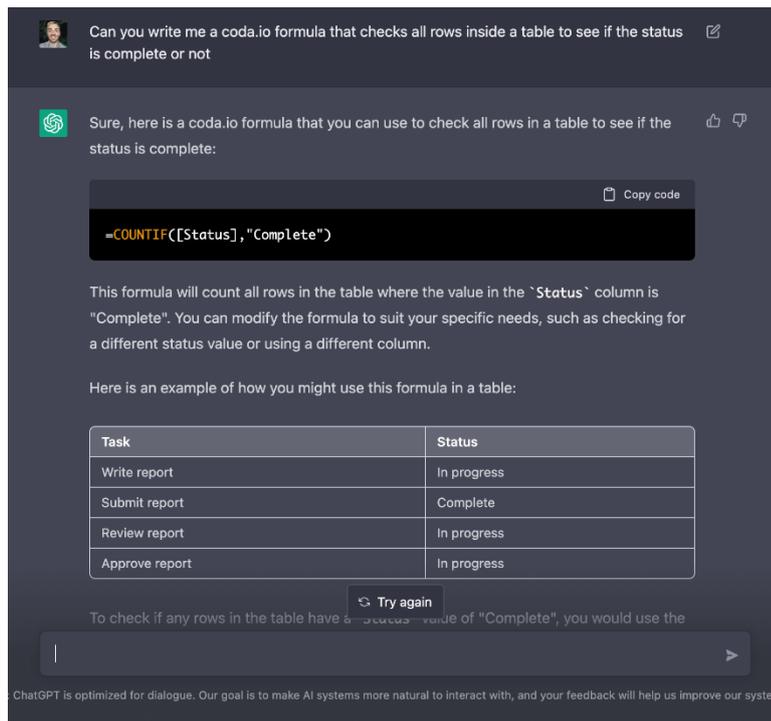
图 8: Customer Assistant Appropriate 数据集 likelihood



资料来源: OpenAI, 中信证券研究部 注: 纵轴为与真实答案的最大相似度

Chat-GPT 的成功, 是在前期大量坚实的工作基础上实现的, 不是横空出世的技术跨越。这些进步主要来自于数学层面的优化带来的结果匹配精准度提高, 而非非算法真正为 AI 带来了创造性与完整的逻辑性, 也不是能够从训练的数据中学习到新的知识。它在“解锁”(unlock)和挖掘 GPT3 学到的海量数据中的知识和能力, 但这些仅通过快速的 In-context 的方式较难获得。InstructGPT 找到了一种面向主观任务来挖掘 GPT3 强大语言能力的方式。因此从这样的底层技术逻辑出发, 我们能迅速找到中短期内适合 Chat-GPT 的产业化方向: 一个真正全方位的智能内容生成助手。

图 9: Chat-GPT 的代码生成



资料来源: ChatGPT

■ 产业影响：成本下降+高质量数据加速底层应用落地

回顾过去近十年间 AI 产业链的发展史，我们可以简单分为以下几个阶段：

1. 小模型（2015 年前）

AI 擅长特定领域的分析任务（如语义理解等），但通用型任务的完成情况很差。硬件算力不够导致重新训练成本过高；数据来源过于稀少难以提升到更高精确度，整体表达能力与人类相差较远。

2. 大模型（2015-2022 年）

Transformer 模型的出现，使得文字、语音、图像识别、语言理解等领域达到了超越人类的水平，但同时也极大得增加了模型的体积，只有拥有强大算力支撑的科技巨头才有能力训练 Transformer 模型。

欢迎访问：电子书学习和下载网站 (<https://www.shgis.com>)

文档名称：从Chat_GPT看生成式人工智能AIGC产业机遇与落地场景.pdf

请登录 <https://shgis.com/post/1775.html> 下载完整文档。

手机端请扫码查看：

