

巨亏，微软 GitHub Copilot 被算了笔账：不赚钱，还要每月倒贴单用户 20 美元

CSDN · 2023-10-11 17:54

关注

你会为大模型产品买单吗？

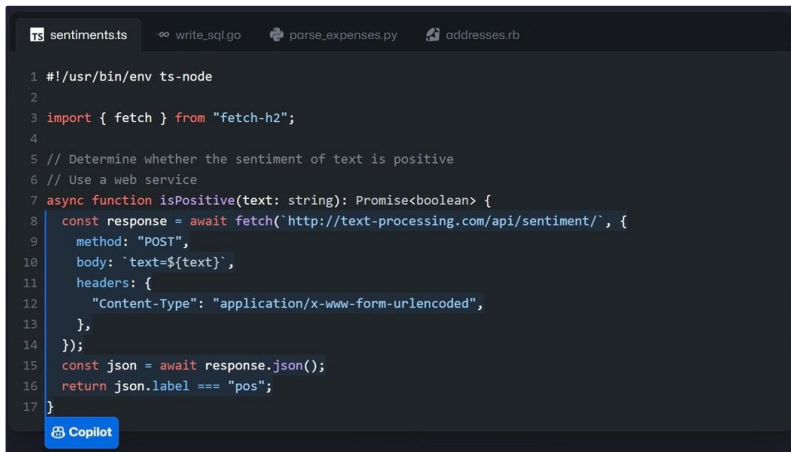
大模型无论是训练还是使用，都比较“烧钱”，只是其背后的成本究竟高到何处？已经推出大模型商用产品的公司到底有没有赚到钱？

事实上，即使微软、亚马逊、Adobe 这些大厂，距离盈利之路还有很远！

同时，使用这些大模型工具的人越多，相关企业需要支付基础设施的费用就越高，正因如此，贴本的买卖也开始了。

01 微软因推出 GitHub Copilot，每月倒贴每位用户 20 美元

继今年 5 月有媒体曝出因去年开发出 ChatGPT，OpenAI 亏损 5.4 亿美元之后，《华尔街日报》最新报道，对于目前在 AI 大模型商业应用维度走在最前沿的微软而言，它在推出的帮助程序员创建、修复和翻译代码服务的 GitHub Copilot 且拥有 150 万用户的基础上，向使用者收取每月 10 美元亦或者是每年 100 美元的费用之际，还贴了不少了钱。



有知情人士透露，在今年年初，微软面向每位用户平均每月损失逾 20 美元，部分用户损失最高逾 80 美元。

此报道一出，引发了不少用户热议，本以为入局大模型是一笔赚钱的买卖，没想到它的成本费用如此高。

02 GPT-4、ChatGPT 为何会成为“吞金兽”？

论及收费 10 美元每月的 GitHub Copilot，其实费用也不算低，那为何还是无法让微软获得营收？

实则，从多个维度来看，这类大模型的运营成本之所以如此高，并非没有理由：

计算资源成本

运行一个大语言处理模型，如 ChatGPT，需要大量的计算资源，包括 GPU 或 CPU。这些资源的租用和维护费用通常占据了大部分运营成本。

此前在与 OpenAI 合作时，微软曾在博客上透露，最初为支持 OpenAI 训练大模型，微软开发了一套新的 Azure 人工智能超级计算技术，也在 Azure 中建立超级计算资源，这些资源的设计和专用性使 OpenAI 能够训练一套日益强大的 AI 模型。

为了训练出这套模型，微软在其基础设施中使用了数以千计的英伟达人工智能优化 GPU，它们被连



作者有点忙，还没写简介

I 发表文章933篇

最近内容

折射OpenAI新一年技术路线图，透视Sam Altman的12个愿望清单

23小时前

“所有代码，都是技术债务”

昨天

历经 64 次更新，开发者倡议“产品别在周五上线”，惹争议后，其 AI 工具意外爆火，...

2023-12-22

阅读更多内容，狠戳这里

下一篇

人工智能和编程的终章

LLM 会对计算机科学带来怎样的影响？


2023-10-11

对此，促成微软和 OpenAI 合作的关键人物——负责战略合作伙伴关系的微软高级主管 Phil Waymouth 表示，OpenAI 训练其模型所需的云计算基础设施的规模是前所未有的，比业内任何人试图建立的网络 GPU 集群都要大得多。后来据彭博社报道，微软在该项目上已经花费了数亿美元。

此外，还有博主根据论文，推算了一下开源大模型的费用：

当训练一个 65B 参数的模型时，我们的代码在 2048 A100 GPU 和 80GB 的内存上处理大约 380 个 token /秒/GPU。这意味着在我们包含 1.4T 标记的数据集上进行训练大约需要 21 天。

2048 个 GPU * 21*24 * 1\$ ~ 100w刀，这还是确定数据集和参数后一次的训练成本



BohuTANG

23-5-4 22:35 发布于 中国台湾 来自 微博轻享版

LLaMa模型花了多少钱呢？他们在论文中提到：“在训练 65B 参数模型时，我们的代码在具有 80GB RAM 的 2048 A100 GPU 上处理大约 380 个token/秒/GPU。这意味着对包含 1.4T token的数据集进行训练大约需要 21 天。”
2048 个 GPU * 21*24 * 1\$ ~ 100w刀，这还是确定数据集和参数后一次的训练成本😂

+关注

数据成本

训练一个类 ChatGPT 模型需要大规模的文本数据集，这些数据集的采集、准备和维护都需要资金。同时，数据的版权和许可成本也可能增加。

早些时候，为了防止数据被“白嫖”，国外诸多问答社区、社交平台纷纷加入数据收费的队伍中，如 Reddit 对每 5000 万次 API 请求收费 1.2 万美元；Twitter 推出最便宜的套餐是每月需支付 4.2 万美元，用户可以访问 5000 万条推文；Stack Overflow 也曾一度被曝要向 AI 巨头收取训练数据费用.....

工程师和研究人员成本

开发、维护和改进 ChatGPT 需要大量的工程师和研究人员。他们的薪资、福利和其他成本会对运营成本产生影响。

不论科技巨头微软，只看 OpenAI。根据 OpenAI 之前发布的官方招聘信息显示，普通的 ChatGPT 软件工程师、机器学习研究科学家的薪资水平在 20 万美元 - 37 万美元（约 138.2 万-255.7 万元人民币）。主管的级别的薪资会更高一些，譬如 ChatGPT 移动端工程主管的薪酬在 30 万美元至 50 万美元（约 207.4 万-345.6 万元人民币）。

安全和隐私

在保护用户隐私和确保系统安全方面的额外投入可能会导致运营成本的增加。这包括监控系统，识别滥用和违规内容，以及实施安全措施。

能源成本

运行庞大的数据中心来支持 ChatGPT 所需的大量服务器和冷却设备，以及相关的能源成本也是一个重要因素。

不久前，微软在发布了一份环境报告，其中透露，2022 年，公司用水量同比去年激增了 34%，较前几年大幅上涨，相当于超过 2500 个奥林匹克规格的游泳池水量，外部研究人员认为这与微软的人工智能研究有关。

维护和更新

模型需要定期维护和更新，以确保其性能、准确性和可用性。这可能需要定期的软件更新、改进和监控。

法规合规成本

符合各种法规和规定，如数据保护法规（如GDPR）和其他监管要求，也可能需要额外的成本，包括法律顾问和合规团队的费用。

性能优化和扩展

不断提高 ChatGPT 的性能和可扩展性可能需要投资于研究和开发，以应对更大的用户需求。

总之，ChatGPT 等大模型的高运营成本是由多个因素共同作用造成的，包括硬件、数据、人力资源、安全、隐私和法规合规等多个方面。这些成本是维持系统高质量运行和用户满意度所必需的，同时也反映了大规模 AI 系统的复杂性。

03 “快速找补”的科技大厂们

有网友认为，「只有亏本经营才能获得迅速市场份额」。

显然，这仅仅是一种短期策略。

为了让大模型更长远地“存活”，微软等入局大模型的公司，已经想尽办法开始“补坑”。

其中，最为值得关注的是，微软正在为其下一次人工智能软件升级提供更高的价格。除了按月定期收取 Microsoft 365 订阅费之外，该公司还将针对人工智能版本每月额外收取 30 美元，付费用户可以享受 AI 驱动的诸多功能，比如快速写邮件、一键把 Word 变成 PPT、快速处理 Excel 图表等。

当然，也并不是微软一家这样做，据悉 Google 也为类似的 Duet AI 产品收取相同的额外费用。

除了提高价格之外，知名外媒 The Information 独家爆料称，微软正在努力开发自己内部 AI 芯片，将其用在数据中心，并推动 PC 行业采用所谓的 NPU。据悉，微软有计划在下个月举行的年度开发者大会上，推出首款人工智能芯片，欲摆脱英伟达芯片带来的高昂费用。

还有一些特殊方式正在入局 AI 大模型的企业中展开，如微软在考虑为 Bing 搜索引擎使用性能较低、成本更低的 AI 工具，还和在开源大模型领域撑起半边天的 Meta 牵手合作了；Zoom 结合了多种 AI 模型，开发了一个更为经济、简单的 AI 来降低成本；Adobe 采取积分制度，对每个月使用量设置了上限，并根据使用情况收费。

整体而言，如果计算成本下降，GitHub Copilot、ChatGPT 和其他人工智能助手的盈利前景将会改变。

04 你会为多少钱的 AI 工具买单？

不过，有网友认为：

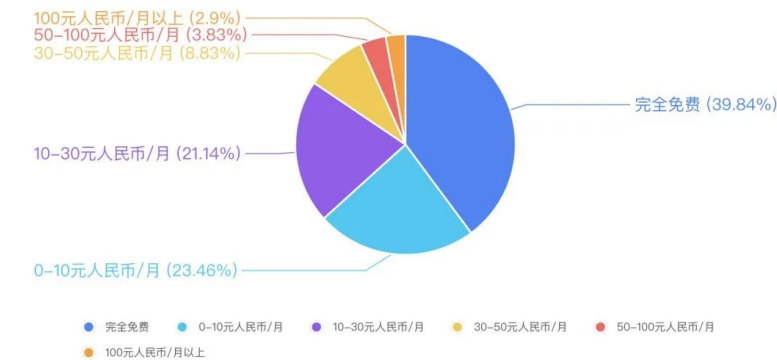
他们真的没有赔钱。用户的体验、观察用户进程、了解用户系统，用户支付微软 10 美元来训练他们的新 AI。10 美元的费用淘汰了没有认真承诺的潜在用户。

也有人表示：

我有 Github Copilot 的企业访问权限，但与免费版的 ChatGPT 相比，它的代码生成能力出奇地差（我主要用它来生成 C 语言的简单函数）。它总是生成不完整的函数，而且似乎无法增加文本生成窗口，甚至无法跟进。此外，我只能通过 VS Code 使用它，而且我也不知道如何使用他们的聊天界面插件。

总的来说，它给人的感觉就是一个不完整的产品。如果能有一个 Web 聊天界面，我就会很高兴。如果他们把服务做得更好一些从而来提升价格，想必也会有不少程序员为此买单。

那么，就目前的现状而言，近日 CSDN 也围绕 AI 开发者生态的最新现状发起了一份调查问卷，根据中期收集到的数据显示，有 39% 的开发者希望代码生成工具完全免费。付费能力在每月 0-30 元人民币区间的开发者，占调查样本的 44%。



对此，你是否愿意为 AI 工具买单？你认为什么样的收费政策是合理的？

本文来自微信公众号“CSDN”（ID:CSDNnews），作者：CSDN，36氪经授权发布。

该文观点仅代表作者本人，36氪平台仅提供信息存储空间服务。



好文章，需要你的鼓励

// 参与评论

评论千万条，友善第一条

[登录后参与讨论](#)

0/1000

[提交评论](#)

// 你可能也喜欢这些文章



Pika 1.0 全面开放测试，打响AI视频圣诞大战，网友脑洞大开一...



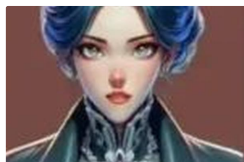
AI如何改变通信？一文了解5G与AI融合的必要性和融合方式及...



OpenAI微软被诉侵权，大模型时代下的版权问题该如何处理？



智源张宏江对话谷歌卢一峰：什么才是大模型「皇冠上的明...



假扮卧底，骗AI泄露代码拯救人类？斯坦福小哥用GPT-4开发...



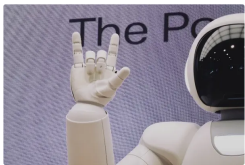
AI主播剑指新闻播报版TikTok，「下一代媒体人」如何应对AI...



Sam Altman的AB面：曾三次被公司“劝退”，硅谷之王如何练成



2024年的生成式AI，到底该投什么？



人工智能：未来社交关系的变革者

关于36氪

- 城市合作
- 商务合作
- 寻求报道
- 关于我们
- 我要入驻
- 联系我们
- 投资者关系
- 加入我们

网络谣言信息举报入口

热门推荐

- 热门资讯
- 热门产品
- 文章标签
- 快讯标签

合作伙伴

- 阿里云
- 火山引擎
- GAUDI 推理
- 个推
- 星球日报
- 鲸准
- 集空网联合办公
- 富途牛牛
- 36Kr 企服点评
- 人人都是产品经理
- 翎氮

36氪APP下载



iOS & Android



本站由 阿里云 提供计算与安全服务 违法和不良信息举报电话：010-89650707 举报邮箱：jubao@36kr.com 网上有害信息举报

© 2011~2023 北京多氪信息科技有限公司 | 京ICP备12031756号-6 | 京ICP证150143号 | 京公网安备11010502036099号



CSDN

特邀作者



5



0

