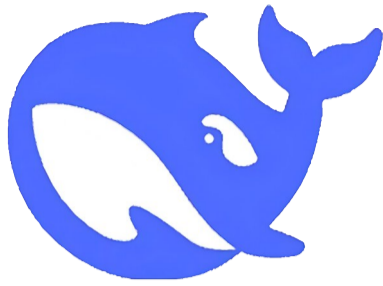


爆火的 DeepSeek 你用了没



DeepSeek 爆火

DeepSeek, 全称杭州深度求索人工智能基础技术研究有限公司, 成立于2023年7月17日, 是一家创新型科技公司, 专注于开发先进的大语言模型和相关技术。

在硅谷, DeepSeek 很早就被称作“来自东方的神秘力量”。2024年底, DeepSeek 发布了新一代大语言模型 V3。2025年1月20日, DeepSeek 发布其最新开源模型 R1, 再次引发全球人工智能领域关注。R1 模型在技术上实现了重要突破——用纯深度学习的方法让 AI 自发涌现出推理能力。这款人工智能助手一经发布就在全球范围内掀起了一股热潮, 在中国区和美国区苹果 AppStore 免费榜均位列首位, 超越了包括 ChatGPT 在内的多个知名 AI 产品。目前, 阿里云、百度智能云、华为云、腾讯云国内四大云巨头都已正式支持 DeepSeek。此前, 海外的亚马逊 AWS、微软 Azure 等云巨头也已官宣支持。

DeepSeek 是什么

DeepSeek 是一款由中国的深度求索人工智能公司研发的大型语言模型, 拥有强大的自然语言处理能力, 能够理解并回答问题, 还能辅助写代码、整理资料 and 解决复杂的数学问题。

与 OpenAI 开发的 ChatGPT 相比, DeepSeek 不仅率先实现了媲美 OpenAI-o1 模型的效果, 还大幅降低了推理模型的成本。其新模型 DeepSeek-R1 以十分之一的成本达到了 GPT-o1 级别的表现, 引发海外 AI 圈的广泛讨论。据报道, DeepSeek 开发的模型在多项测试中的表现都优于 OpenAI, 且设计成本不到 600 万美元, 堪称“小力出奇迹”。

DeepSeek 代表了 AI 的未来方向——更懂人话, 更会思考, 它不仅“脑子快”, 还擅长“个性化发挥”, 为你做到“全方面考虑”, 被网友称为“全能型 AI 助手”。

DeepSeek 再获「朋友圈」

阿里云、百度智能云 2 月 3 日发文宣布, 正式接入 DeepSeek-V3、DeepSeek-R1 模型。加之华为云、腾讯云, 目前国内四大云巨头都已正式支持 DeepSeek。此前, 海外的亚马逊 AWS、微软 Azure 等云巨头也已官宣支持。

多家国产芯片厂商也在积极响应。天数智芯、摩尔线程 2 月 4 日接连发文宣布支持 DeepSeek 模型。GiteeAI 联合沐曦 2 月 2 日首发全套 DeepSeek-R1 千问蒸馏模型, 实现了从芯片到平台, 从算力到模型, 全国产研发、中国制造。

业内人士认为, DeepSeek 以创新的技术路线吸引了全球科技巨头及优秀独角兽企业的关注和支持, 说明国产科技在开放创新的开源环境里具备强大的竞争力, 中国本土科技企业的价值理应在被重估。

2 月 2 日—4 日, 腾讯云、百度智能云、阿里云等先后接入 DeepSeek 相关大模型, 而大厂在春节假期为深度求索公司加班, 只是 DeepSeek 爆火的一个证据。英伟达市值蒸发创纪录、DeepSeek 卡顿、榜单第一……与深度求索直接间接相关的新闻, 在春节期间不断刷新。

DeepSeek 是什么? 为何震惊世界?



DeepSeek 将推动 AI 实现平民化

随着 DeepSeek 掀起新一轮 AI 热潮, 不少普通人会问, 未来我们的生活是否会因此而改变?

正如苹果 iPhone 的发明被认为是一项颠覆性的技术, 因为它彻底影响了人们的生活方式, 让千千万万人享受到了技术的红利。在一些业内人士看来, DeepSeek 将推动 AI 实现“平民化”, 而建立在开源的 AI 系统之上, 更多的开发者能够开发出有用的 AI 应用场景, 从而让 AI 走入普通人的生活。

AI 开源初创公司 Hug-gingFace 联合创始人、首席科学家 Thomas Wolf 在日前的公开活动上谈论到 DeepSeek 时表示, 人工智能的下一步将是“应用”。他说:“大家都在谈论性能强大的 AI 大模型, 但需要强调的是, AI 在应用端还有非常长的‘长尾效应’, AI 系统的开源将促进这种‘长尾效应’的发挥。”“很多模型已经开源并且向用户免费提供。”他说道, “下一步将是如何在这些模型之上建立应用, 让模型在应用层面变得有用, 从而能完成现实中的任务, 让每个人能在生活中受益。”

业内人士表示, DeepSeek 规模较小的模型将助推 AI 模型更广泛的部署, 降低人们使用 AI 的门槛。

遭暗中阻击

据媒体报道, 1 月 30 日凌晨, 即农历大年初二, 奇安信 XLab 实验室监测发现, 针对 DeepSeek (深度求索) 线上服务的攻击烈度突然升级, 其攻击指令较 1 月 28 日暴增上百倍。XLab 实验室观察到, 至少有两个僵尸网络参与攻击, 共发起了两波次攻击。

据央视报道, 当地时间 2 月 3 日, 在美国纽约联合国总部, 中国常驻联合国代表傅聪

举行记者会。当被问到中国人工智能企业深度求索 (DeepSeek) 发布的模型及中美人工智能合作问题时, 傅聪表示:“从华为到 TikTok, 再到 DeepSeek, 美国还想禁多少?”“永远不要低估中国科研人员的聪明才智。DeepSeek 引发全球轰动和一些人的焦虑恐慌, 说明技术遏制和技术限制无法奏效, 这是全世界、特别是美国需要学习的一课。”

攻击指令激增 100 多倍

XLab 实验室通过对 DeepSeek 持续近 1 个月的监测发现: 攻击模式从最初的易被清洗的放大攻击, 升级至 1 月 28 日的 HTTP 代理攻击 (应用层攻击, 防御难度提升), 现阶段已演变为以僵尸网络为主。攻击者使用多种攻击技术和手段, 持续攻击 DeepSeek。

1 月 30 日凌晨, XLab 观察到两个 Mirai 变种僵尸网络参与攻击, 此次攻击共涉及 16 个 C2 服务器的 118 个 C2 端口, 分为两个波次, 分别为凌晨 1 时和凌晨 2 时。

“僵尸网络的加入, 标志着职业打手已经开始下场, 这说明 DeepSeek 面对的攻击方式一直在持续进化和复杂化, 防御难度不断增加, 网络安全形势愈发复杂严峻。”奇安信 XLab 实验室安全专家表示。

僵尸网络是由攻击者通过恶意软件感染并控制的设备网络, 这些设备被称为“僵尸”或“机器人”。攻击者通过命令与控制 (C&C) 服务器向这些设备发送指令, 执行各种任务, 例如向目标服务器同时发起 DDoS 攻击, 持续增加攻击规模和强度, 耗尽目标服务器的网络带宽和系统资源, 使其无法响应正常业务, 最终瘫痪或服务中断。

本次采用的两个僵尸网络分别是 HailBot 和 RapperBot, 这两个 Botnet 长年活跃, 攻击目标遍布全球, 专业为他人提供 DDoS 服务。其中, RapperBot 平均每天攻击上百个目标, 高峰期指令上千条, 攻击目标分布在巴西、白俄罗斯、俄罗斯、中国、瑞典等地区。

中国明星产品屡遭暗中阻击

DeepSeek 推出 R1 模型后不久, 就凭借其性价比、开源及推理能力的提升等方面获得了广泛关注。除夕当天, DeepSeek 还推出了新模型, 其中 Janus-Pro-7B 在基准测试中击败了 OpenAI。

DeepSeek 的成功不仅引发了硅谷的震动, 更让华尔街感到紧张。就在 1 月 28 日, 美国芯片巨头英伟达一夜市值蒸发 5900 亿美元, 合 4.3 万亿人民币, 纳斯达克综合指数跌 3.07%, 台积电、博通公司、超微半导体等科技股也遭遇集

体暴跌。美国总统特朗普表示, DeepSeek 的崛起应当为美国企业敲响“警钟”, 美国公司“需要专注于竞争以赢得胜利”。美国众议院首席行政事务官向国会办公室发出通知, 警告国会办公室不要使用中国的人工智能应用 DeepSeek (深度求索) 的服务。

每次中国优秀的明星产品或企业崛起之时, 总会遭到一些境外不法势力的暗中阻击。上一次是《黑神话: 悟空》全球上线后, 遭遇了海外 60 个僵尸网络大规模攻击。

DeepSeek 为何让美国紧张

原因一: 高性价比 冲击美国大模型垄断地位

DeepSeek 大模型训练成本不到 600 万美元, 仅为同性能模型的 5% 到 10%。新模型训练方法大幅度降低了大模型行业的人局门槛, 大规模训练不再是科技巨头的专利。

原因二: 模式创新 带来美国高新技术人员恐慌

在喂养学习数据这一大模型重要环节上, OpenAI 选择了“人海战术”, 堆砌算力、将资源集中在算力, 用海量数据投喂实现能力的提升。而 DeepSeek 相比于“砸资源”选择了另外一种方式。利用算法把数据进行总结和分类, 经过选择性处理之后再输送给大模型, 最大优化算力实现了

成本的降低和模型性能提升。而这引发了不少美国企业技术人员的恐慌, 他们担心自己被质疑技术能力和创新性从而失去工作。

原因三: 进步显著 国产大模型正在厚积薄发

根据中国工业互联网研究院推出的《人工智能大模型年度发展趋势报告》, 与国际顶尖大模型能力相比, 2024 年国内大模型的能力进步显著。从 2023 年第四季度到 2025 年第一季度的测评显示, 国内外大模型能力差距缩小了将近 75%。可以看出, DeepSeek 的出现并不是所谓的“异军突起”, 而是中国国内大模型整体发展的阶段性成果体现。

据《华商报》报道