

2025年2月

7

星期五
农历乙巳年
正月初十

山西科技报

立媒 科经

科技改变世界 创新引领未来

周刊

新闻热线
4159087

国内统一连续出版物号 CN 14-0009 邮发代号 21-17 总第 7821 期 今日 8 版

本刊邮发代投号 21-1016

阅读提示

2024年12月26日,杭州深度求索人工智能基础技术研究有限公司(中文简称深度求索,英文简称 DeepSeek)发布了 DeepSeek 系列模型——DeepSeek-V3,并同步开源。因其在性能上与全球顶尖闭源模型 GPT-4o 和 Claude-3.

5-Sonnet 不相上下,所以迅速引爆 AI 圈,并冲上热搜。

那么,DeepSeek 是什么?它有何功能?如何学习使用?就此,小编就手头掌握的资料,与读者进行科普。

这个人工智能大模型为何“牛”?

——从 DeepSeek-V3 冲上热搜说起

DeepSeek 是什么?

DeepSeek 是一个智能系统。可以理解、分析和生成信息,帮助人们解决问题或完成任务。

DeepSeek 系列模型,核心产品除 DeepSeek-V3 外,还有 DeepSeek LLM、DeepSeek-V2、DeepSeek Coder、DeepSeek Math、DeepSeek-VL 和 DeepSeek-VL2 等。

这些模型不仅完全开源,还支持商用,提供了多种接入方式,如 AI 对话、API 接口和模型下载,方便用户快速集成和使用。

DeepSeek 的三大功能

你可以把 DeepSeek 想象成一个“超级助手”,它能够通过大量的数据和算法,快速找到你需要的答案,或者帮你完成一些复杂的任务。

1. 理解信息:像“读心术”一样懂你

例如,你问 DeepSeek:“我想去一个安静的地方度假,有什么推荐吗?”

DeepSeek 接到你的指令后,它会先理解你的需求——需要找一个安静的地方、能满足度假条件。然后,结合你的开支预算和偏好,推荐一些冷门但风景优美的地方,比如“云南的普者黑风景区”或“浙江的莫干山景区”。它还会告诉你这些地方的特点,比如“普者黑有美丽的湖泊和田园风光,适合放松心情”。

2. 分析信息:像“侦探”一样找答案

例如,你是一名创业者,正在研究“如何提高用户留存率”,请 DeepSeek 帮助你分析。

DeepSeek 会直接分析海量的商业案例、用户行为数据和市场研究报告,提取出关键信息。比如提高用户留存率的三大策略:“优化用户体验、增加互动性、提供个性化服务”,并帮你整理成清晰的行动计划。

3. 生成信息:像“创作者”一样帮你解决问题

例如,你正在准备一场朋友聚会,想设计一个“适合 10 人左右的互动游戏,轻松有趣,时间控制在 30 分钟内”。这时,可以问 DeepSeek。

DeepSeek 接到你的指令后,会能生成几个创意游戏方案。比如:

(1)你画我猜升级版:每人轮流画出抽象的概念。比如,时间旅行。再由其他人猜。



(2)故事接龙:每人说一句话,串联成一个完整的故事,最后大家一起读出来。

(3)快速问答挑战:设置一些搞笑或冷知识问题,看谁答得最快最准。

如果你觉得个游戏太复杂,它还会根据你的反馈调整规则,比如“简化你画我猜的题目,让游戏更轻松”。就像一个创意无限的派对策划师,帮你轻松搞定聚会!

所以说,DeepSeek 不仅仅是一个工具,而是你的“全能型智能生活助理”。

无论是学习、工作还是生活,它都能为你提供强大的支持。

当你迷茫时,它像一盏指路灯,帮你找到方向。

当你忙碌时,它像一个高效助手,替你完成任务。

当你需要灵感时,它像一个创意引擎,为你打开新思路。

DeepSeek-V3:花小钱办大事

在 DeepSeek 系列产品中,DeepSeek V3 完美延续了其产品“使用便宜”的优良传统,一发布就毫无保留地选择完全开源,甚至用 53 页满满当当的论文,将训练细节毫无保留地公之于众。

DeepSeek V3 这匹“大模型黑马”,能和 GPT-4o、Claude 3.5 Sonnet 这些 AI 界的

“武林高手”正面过招,毫不逊色。更让人惊掉下巴的是,它的价格低到超乎想象,仅为 Claude 3.5 Sonnet 的 9%,性价比直接拉满!

训练时长和成本上,DeepSeek V3 更是开启了“降维打击”。整个训练过程仅用不到 280 万个 GPU 小时,而 Llama 3 405B 却要 3080 万 GPU 小时,简直是天壤之别。换算成真金白银,训练 671B 的 DeepSeek V3 成本是 557.6 万美元(约合 4070 万人民币),而训练一个 7B 的 Llama 2 就要 76 万美元(约合 555 万人民币),这差距,就像坐火箭和坐牛车。OpenAI 创始成员 Karpathy 忍不住点赞,Meta 科学家田渊栋也直呼其训练是“黑科技”。

从模型能力来看,它的评测跑分一骑绝尘,不仅把 Qwen2.5-72B 和 Llama-3.1-405B 等开源模型远远甩在身后,还和顶尖闭源模型打得难解难分。实际响应速度更是快得飞起,提升了 3 倍!

这是因为,DeepSeek-V3 通过更先进的 MoE 架构、多技术融合优化、FP8 混合精度训练框架等技术,以及与开源社区合作的方法。所以,在成本较低的情况下,DeepSeek-V3 就训练出文字生成和逻辑推理能力。

(下转 04 版)

>> 导读

富士康太原园区工会 举办送春联活动

山焦汾西中盛煤业

上好复工安全“第一课”

晋能控股煤业集团麻家梁
煤业公司综采二队检修班

当好综采设备 安全“护航兵”

02 版

山西大学校领导向一线 在岗职工送新春祝福

朔州市委副书记、政法委
书记宋红波一行

走访慰问 造血干细胞捐献者

太原市妇幼保健院

召开晋籍医学专家 年终工作经验交流会

03 版

为什么有的人 看电脑屏幕时, 眼睛会不舒服?

05 版



新闻传播更及时 / 新闻内容更生动 / 新闻受众更广泛

主管:山西省科学技术协会 主办:山西科技新闻出版传媒集团 出版:《山西科技报》社有限责任公司 出版单位地址:太原市新建路 18 号 印刷厂:太原日报传媒集团印务有限公司(唐槐路 80 号) 总编辑:王而平

广告经营许可证号:1400004000004 广告/发行热线:4159062 邮箱:sxkj@163.com 邮编:030002 发行方式:邮局 发行单位:山西省邮政报刊发行局 全年定价:380 元 头版责编:李恒灵