

AI 改变生活, DeepSeek 是“催化剂”吗?

最近, DeepSeek 可以说是全球瞩目。作为一种开源人工智能技术, DeepSeek 通过其深度学习模型 DeepSeek-R1 迅速成为全球人工智能(AI)技术的领军者之一。它不仅在数学推理、代码生成、自然语言理解等方面展现了超越现有模型的能力,还通过降低技术的准入门槛,给全球 AI 应用的迅速普及带来无限可能。

DeepSeek 的成功更是凸显了人工智能技术发展的重要趋势——从单一的推理向更高层次的认知能力演进。越来越多的企业和开发者开始关注如何将这种技术应用到日常生活和工作中,尤其是随着 AI 推理成本的逐渐下降, AI 开始走进每个普通人的生活。

突破与普及

DeepSeek-R1 模型的技术架构由超过十亿个参数组成,能够处理复杂的推理任务,其高效的计算能力,使得中小企业和开发者也能负担得起高性能的 AI 应用。与传统的 AI 模型不同, DeepSeek 采用了开源模式,这意味着全球的开发者都可以自由使用、修改和提升这个模型。这一策略不仅推动了全球技术创新,也为广大开发者和企业带来了更多机会。

这一模式与传统的闭源 AI 模型相比,显然更具优势。闭源模型通常依赖商业化盈利模式,技术和创新的门槛较高,而开源 AI 的普及让更多人可以参与其中,推动大模型的“平权”和普惠 AI 的到来。

影响你我

随着 DeepSeek 等模型的普及, AI 正在逐步改变我们的日常生活和工作方式。首先, AI 正在医疗领域发挥重要作用,帮助医生分析医学影像、提供诊断建议。在一些基层医院, AI 可以通过分析症状和病例,协助医生做出更加准确的决策,甚至有可能协助解决医疗资源不足的问题。

其次,在金融领域, AI 能够处理海量数据,预测市场趋势,帮助投资者制定更精准的投资策略。

最后,对于普通消费者来说, AI 也能帮助我们通过智能算法选择最佳的消费方案,如精确推荐个性化的商品或服务,甚至优化生活中家务规划、交通出行、健康管理等方面的决策。

另外, AI 还带来了生产力的提升。比如,在制造业中, AI 已经能够帮助企业优化生产线、减少人工成本。在交通领域,智能交通系统的出现则有助于提高交通通行效率,缓解拥堵问题。

然而, AI 带来的影响不仅仅局限于技术应用,它还会影响到普通人的就业。一些低技能岗位可能面临被替代的风险,尤其是在制造、客服等行业中尤为明显。但也意味着许多人将进入新兴领域,成为 AI 技术的使用者和管理者。如何提升自己的技能,适应 AI 时代的变化将变得更加重要。

是挑战也是机遇

DeepSeek 等开源技术虽然已经取得了显著的成果,但距离实现真正的通



用人工智能 (AGI) 还有一定的距离。AGI 是指能够像人类一样理解和解决各种复杂问题的智能系统。尽管现在的 AI 在某些领域的表现已经足够优秀,但仍然缺乏人类的创造性思维和跨领域的知识迁移能力。

AGI 的挑战远不只是停留在技术层面上,还包括伦理、法律等多方面的考虑。随着 AI 的普及,数据隐私、知识产权、技术监管等问题也变得愈加重要。如何平衡技术创新与治理,将是未来 AI 发展过程中需要解决的关键问题。

随着计算资源的不断优化和 AI 服

务成本的下降,更多企业和个人开发者能够负担得起 AI 技术。此外, AI 将在推理和认知能力上不断提升。未来, AI 不仅能处理文本,还能更好地理解图像、音频等多模态信息,进一步增强与人的交互能力。在人机协作的场景下, AI 将不仅是工具,而是与人类共同解决问题的合作伙伴。

可以预见,以 DeepSeek 为代表 AI 技术正在深刻地改变我们的世界。对于普通人而言,学习如何与 AI “共舞”,应是迎接这一新机遇的关键。

据《科普时报》报道

“AI 医生”来了,未来将应用于哪些场景?



近日,多家医院宣布,将 AI 医疗落地到医院诊疗场景中。AI 技术将为医疗领域带来哪些变化?“AI 医生”能给人看病治病吗?

日前,全国首个“AI 儿科医生”在国家儿童医学中心北京儿童医院正式上线应用,这名存在于虚拟世界里的“儿科医生”,承担的角色主要有两个:

一个是专家的临床科助理,帮助医生快速获取最新科研成果和权威指南;

另外一个则是辅助医生进行疑难罕见病的诊断和治疗。

就在本月 20 日,多位身穿白大褂的知名专家,和身穿黑色工装服操作下的“AI 儿科医生”,共同为 10 名患儿进行疑难病多学科会诊。

就在专家们讨论的同时,桌子的一边,工程师将患者的病历资料输入大模型中,“AI 儿科医生”也给出了治疗方

案:手术加化疗。记者观察到,10 名患儿的诊断中, AI 儿科医生给出的建议与专家组会诊结果吻合度较高,有时 AI 儿科医生还会通过患者病例报告里的某个单项指标,给治疗方案增加特别提示。

不同于通用 AI 大模型,医学专业领域的大模型需要更多维度的精准数据。为此,国家儿童医学中心将 300 多位专家的临床经验及数十年高质量病历数据,整合进大模型中,从而把“AI 儿科医生”训练得更专业、准确。

国家儿童医学中心主任北京儿童医院院长倪鑫:我们在研发儿科 AI 医生初衷,当时提出“双医生制”。我们所有的结果跟 AI 相互提醒, AI 对于整个医生团队就是辅助和朋友。

人工智能在医疗领域还有一个功能是预判,相当于医学检测多了一把“尺

子”。在北京协和医院,工作人员正在为陈女士做三维步态评估。

去年以来,陈女士左侧肢体发抖、发僵,走路越来越困难,医生怀疑她是神经系统疾病。高速摄像机记录下了陈女士的运动过程,然后交给人工智能系统去评估。

一直以来,神经系统疾病在临床上不同医生的评估结果存在差异,容易出现误诊、漏诊。“面向神经系统疾病预警的智能人机交互关键技术”,通过多维评估体系,助力阿尔茨海默病、帕金森等神经系统疾病的早期诊断。

北京协和医院神经内科医生 苏宁:这个病人的评分是 63 分,有可能是在认知障碍的早期。

截至目前,这套系统已完成 10 多万例神经功能检测,为临床精准诊疗提供参考依据。

北京协和医院神经内科主任 朱以诚:比如说我一年前看了一个帕金森的病人,我不记得他一年前在我面前是什么样子了,我没有一个量化值,但机器都记下来了。它作为一个长期随访的测量工具是非常有用的。

“AI 医生”来了 如何辅助医疗?

如今在康复医疗领域,人工智能技术已经实现多方面的应用。在杭州,这个外骨骼机器人,正在帮助残障人士进行康复训练。朵朵十七岁时,一次意外摔伤令她下肢瘫痪。经过康复训练,朵朵的身体一天天好转。

心理咨询和心理教育领域也正在牵手人工智能。北京大学第六医院的“北小六”机器人已经在多地上岗,它能先和用户进行几轮 AI 对话,判断用户在认

知、情绪、行为、人际关系等领域存在的问题,并评估问题的严重程度并给出相应建议,或提示用户到医院精神科接受进一步治疗。

在一些学校, AI 正助力教师成为学生知心伙伴。在北京市宏志中学的学生心理驿站,有两个不需要预约的“心语小屋”,里面的 AI 大模型学习了近 40 万篇心理学期刊文献和来自医院、专业机构的 100 万个心理辅导对话案例,帮助学生大胆说出心里话。

北京市宏志中学心理教师周冬梅:学生可以输入自己的账号、密码,私密性就得到保障。孩子就能够去畅快地吐露自己的心声。

湖南省长沙市湘江新区莲花镇中心小学是一所镇小学,1200 多名学生中不少是留守儿童。心理教师通过学生与 AI 的互动测试,动态掌握学生的心理变化。

湖南省长沙市湘江新区莲花镇中心小学校长周燕:它可以有效解决我们很多乡镇学校在心理健康师资源上的不足。

AI+医疗 未来还将应用于哪些场景?

去年 11 月,国家卫生健康委等三部门发布《卫生健康行业人工智能应用场景参考指引》,给出了 84 个典型应用场景。

高频词为“智能”和“辅助”。“智能”体现出 AI 在技术层面的价值,而“辅助”则强调了 AI 在医疗领域承担的工具性角色。

有医生表示, AI 的确可以在一些场景,一些环节辅助医疗,但医生面对复杂病情的综合判断和经验积累以及面对的人文关怀,是 AI 无法替代的。

湖南郴州市儿童医院儿童健康研究所执行所长肖能:它可以总结所有目前已发表的文献,所以我们可以给患者提供适合的诊断方案。

四川省医学科学院四川省人民医院副院长傅启华:患者只需打开我们医院的微信公众号,用语音或者文字询问“胃疼怎么办”或“头疼应该挂哪个科室”, AI 就能迅速给出贴心的建议和就医指南。

湖南怀化第二人民医院脊柱关节科康复医学科护士长 康丹:一些自动化的任务,如智能导诊、药物分配等重复性和机械性的护理任务,正在被自动化系统取代。但是也需要特别警惕伦理、护理的数据安全和患者的隐私保护等问题。

北京协和医院罕见病医学科主任沈敏:现在有一个词语叫“AI 幻觉”,就是你拿到的这些信息到底是真的还是假的?我们要去甄别,因为它也是通过海量的数据分析。那些海量数据的里面,也许混杂了一些假消息。而且我们不是冰冷的机器,我们还要更多体现人文的决策。

据《央视新闻客户端》报道

当前,基于生成式人工智能(AIGC)伪造技术的电信网络诈骗手段不断升级,呈现范围扩大化、套路多样化、形式隐蔽化等特点,危害社会安全。科研人员正在开发更先进的人工智能工具和技术,识别并阻止人工智能滥用行为。

“不管是以往的 PS 篡改识别,还是如今的 AIGC 鉴别,都属于人工智能鉴别的被动检测方式。”在第十三届中国创新创业大赛创新挑战赛(浙江)暨 2024 年浙江省技术需求“揭榜挂帅”大赛总决赛上,杭州电子科技大学网络空间安全学院副教授、浙江-法国数字媒体取证联合实验室负责人乔通展示了其团队研发的“面向电诈场景的音视频鉴伪溯源技术”,这种技术变“被动检测”为“主动防御”,有望为人工智能鉴别提供新思路。

特殊水印实现“事前防御”

上网搜索“AI 换脸”,相关应用程序和影音图像作品比比皆是。同时,深度人脸伪造带来的安全威胁也逐渐显现。

香港警方曾披露一起人工智能“多人换脸”诈骗案:不法分子利用人工智能深度伪造技术,通过某跨国公司社交平台上的资料,伪造多位高层管理人员的形象和声音,在视频会议中进行冒充,诱骗该公司香港分公司财务职员进行转账操作,涉案金额达 2 亿港元。

此前,针对人脸伪造,人们往往采用“事后鉴别”的被动检测、被动取证方法,但这种方法难以及时、有效遏制伪造人脸图像的传播。主动防御方法由于“事前防御”的属性,得到越来越多学者关注。

然而,目前大多数主动防御方法只能针对单一或特定的几种伪造模型进行防御,面对日益多样化和不断升级的人脸伪造技术,会显得力不从心。

为此,乔通团队开发出了一种可扩展通用对抗性(SUA)水印,将其嵌入在需要保护的音频、图像或语音素材中。当素材被用于人工智能加工,就会触发 SUA 水印的“引信”,直接破坏生成后的内容。

这一技术路线此前已有所应用,但局限性较大。乔通介绍,以往单个水印只能针对一种模型,一段素材可能要嵌入多个水印才能生效,而 SUA 水印则无须基于所有伪造模型进行重新训练。大量实验证明, SUA 水印不仅能节省算力,还能有效扩展主动防御范围,显著提高主动防御效果,能大大提高鉴别效率。

“火眼金睛”进行人脸识别

2024 年 9 月,杭州公安通报了一起利用人工智能技术侵犯公民个人信息案件:有不法分子使用人工智能换脸大模型突破头部平台人脸识别,窃取用户隐私信息。

“登录银行或者其他 App 时,用户会被要求做摇头、眨眼的动作,即进行人脸活体检测。”乔通说,不法分子通过黑色产业链,可以获取目标对象的人脸照片、音频素材,伪造多模态媒体素材,然后实施诈骗。

为保证平台安全,乔通建议,进行人脸识别前应再加一道无形的核验环节。据了解,针对目前市场上主流的人工智能伪造工具,乔通团队开发的人工智能鉴别系统,能秒速级识别出人工智能图像及视频,并给出相应分析报告,准确率 95% 以上。

“相比大型银行,地方中小银行对防伪技术的投入相对不足。”乔通说,团队目前在与地方公安、小型金融机构等进行合作,研发定制化技术,在活体检测环节为相关平台装上“火眼金睛”,判定屏幕那端“是人还是妖”。

在乔通看来,人工智能鉴别和深度伪造是此消彼长的关系,需要更多机构、企业参与,推动人工智能鉴别技术突破和应用落地,共同搭建人工智能时代的网络安全防火墙。“我们团队正积极对接企业,探索商业化模式。”乔通说,“要筑牢人工智能安全底座,让反诈举措更加柔性。”

据《科技日报》报道



人工增雨还是人工降雨?

提到气象服务,想必很多人最先想到的就是人工影响天气,但对其中更专业的科学问题并不是很了解,比如到底是人工增雨还是人工降雨?什么条件才能进行人工增雨?

想要了解这两个问题,首先要了解人工影响天气业务是指为避免和减轻气象灾害,合理利用气候资源,在适当条件下通过科技手段对局部大气的物理、化学过程进行人工影响,实现增雨、防雹、消雨、消雾、森林防(灭)火、消除公共污染等的业务活动。然而对于乌兰察布地区来说由于地处内蒙古自治区中部,属典型的中温带半干旱大陆性季风气候,所以人工增雨是本地主要的人工影响天气业务。

但公众并不理解到底是人工增雨还是人工降雨,一字之差概念却完全不一样。人工增雨作业并不是凭空造雨,而是在有降水条件时人为将催化剂送到云层中促进小水滴快速集结形成大水滴像“卤水点豆腐”一样把云中的水汽“点化”成雨滴使本来不会产生的降水得以产生,从而达到增加降水的效果。同时人工增雨并不是想增就能增,而是需要一定的气象条件,即云中水汽充沛、有一定的上升气流,有足够多的凝结核或冰核,才具有成熟的增雨条件。

如今,人工影响天气给人们的生活、社会的发展都带来了非常显著的好处;未来,气象部门会进一步改进和完善人工影响天气技术,使之在保护农业生产、缓解水资源短缺、防灾减灾、保护生态等方面发挥更大的作用。

(乌兰察布市气象局 刘薇)

『被动检测』变为『主动防御』之鉴伪技术开始反制『换脸』诈骗