



聚焦 ChatGPT

国内研发进展如何? 中国版 ChatGPT 下月亮相

ChatGPT 大受欢迎,其重要幕后功臣是被业内称为预训练模型的技术。面对 ChatGPT 的爆火,国内科技机构和企业纷纷“亮剑”。据最新消息,中国版 ChatGPT 将于下个月与公众见面。



1 参数规模从几百万狂飙至几千亿

“目前,预训练技术是人工智能研究的重要突破口。”阿里研究院未来技术中心负责人苏中解读,传统的研究方法中,标注成本一直是阻碍人工智能算法推向更大数据集的障碍,而预训练技术不依赖数据标注,就可以训练出一个大规模深度学习模型。

记者了解到,由于不需要数据标注,预训练模型往往可以使用更大的数据集,因而可以选择更大的模型规模——这就催生了预训练大模型。

“与以往的人工神经网络相比,预训练大模型最大的不同在于它规模足够大,深度学习网络的层数多、连接多、参数多。”IDEA 研究院(粤港澳大湾区数字经济研究院)认知计算与自然语言研究中心讲席科学家张家兴接受记者采访时介绍,2012年前后的深度学习网络只有几百万参数;2018年前后主流的预训练模型达到1亿参数;目前被证明非常有

效的大规模预训练模型已有几千亿参数,短短几年时间提升了几千倍。

量变引起质变。苏中分析,在对预训练模型各种不同的技术评测中,算法性能展示了一个规律:数据规模越大、预训练模型参数越多,算法输出精度往往也越高。

除了规模大以外,张家兴谈到,目前的预训练大模型由于采用了新的结构模型,非常适合并行训练。同时,在训练方式上,人工神经网络的预训练任务类似“解题”,比如把一句话的某个词或前半句去掉,让预训练模型利用海量大数据对句子进行恢复,这样训练下去,预训练模型的“解题”能力就会越来越好。而且,在大语言模型的基础上,引入指令学习,可以远超模型微调带来的收益。在此基础上,人类的反馈可以帮助模型更好地理解人类意图和接收信息,人机对话也就越来越流畅。

2 国内科技机构已布局相关领域

据外媒报道,多家国外科技巨头正重金投入预训练模型和人工智能生成内容(AIGC)赛道。记者了解到,国内已有不少科技机构和企业布局相关领域。

张家兴告诉记者,IDEA 研究院认知计算与自然语言研究中心在此方向上布局已久,该中心拥有国内最大的开源模型体系——封神榜,已开源的98个模型正在全力支持中国 AIGC 的发展,探索研发适配各个领域的个性化可定制模型,支持全部自然语言处理(NLP)任务。

“很高兴能在深度学习高速发展的第十年节点上,看到 AIGC 和 ChatGPT 的出现,今天所有这些自然语言处理任务正在被统一到 ChatGPT 模式下,我们也正在整合之前的模型和技术,推出我们的类 ChatGPT 模型。”张家兴说。

记者从百度获悉,百度已投入开发类似 ChatGPT 的相关技术,该项目名字确定为文心一言,英文名 ERNIE Bot,3

月份完成内测后面向公众开放。目前,文心一言在做上线前的冲刺。

“ChatGPT 主要涉及自然语言处理相关技术,属于认知智能领域的应用之一,科大讯飞在该方向的技术和应用方面具备长期深厚的积累。”2月8日晚,科大讯飞在投资者互动平台回应称,公司已面向认知智能领域陆续开源了6大类、超过40个通用领域的系列中文预训练语言模型,并于去年12月启动生成式预训练大模型任务攻关,人工智能学习机将成为该项技术率先落地的产品,今年5月将进行产品级发布。

此外,360公司2月7日在投资者互动平台表示,该公司也计划尽快推出类 ChatGPT 技术的原型版产品。“AIGC 技术除了在搜索引擎的应用之外,还可以辅助数字安全能力的提升。”360公司称。

综合央视、北青网、科技日报等媒体报道

这些“ChatGPT”是假的 小心被骗!

随着聊天机器人 ChatGPT 火爆全球,国内也涌现了一批名字中包含“ChatGPT”的微信公众号、小程序产品。

国内这些产品是真的吗?与国外的 ChatGPT 有什么联系?

用 ChatGPT 官方图标当头像 免费试用几次后开始收费

在微信上搜索,可以发现一系列与 ChatGPT“沾亲带故”的产品。

这些产品的服务方式大多是先免费试用,一旦免费次数用尽就开始收取费用。

以“ChatGPT 在线”为例,它为用户提供4次免费对话额度,之后继续使用需充值,充值额度分别为9.99元/20次(三个月有效)、99.99元/1300次(半年有效)、199.99元/3000次(一年有效)、999.99元/无限次(一年内有效)。

而另一款类似的服务“GPT 深蓝”也显示有199元月度会员、399元季度会员、999元年度会员的 ChatVIP 充值机制,页面甚至还有“加入代理赚钱”的选项。

并未在中国大陆提供有关服务 “山寨版”涉嫌商标、名称侵权

目前,ChatGPT 的开发者、人工智能研究机构 OpenAI 并没有在中国大陆提供有关 ChatGPT 的服务,也没有相应的公司运营。

对于使用 OpenAI、ChatGPT 的名称、商标等行为,北京周泰律师事务所高级律师储江表示,此类产品在网络上的活动涉及一系列法律问题,存在违反《商标法》《反不正当竞争法》《消费者权益保护法》《网络安全法》甚至《刑法》相关规定的可能。

储江称,ChatGPT 是由 OpenAI 推出的产品,由 OpenAI 享有其一切的财产性权益,这其中就包括作为产品/服务来源标识的“ChatGPT”名称。

但储江介绍,我国《商标法》主要保护的是注册商标权。“自然人、法人或者其他组织在生产经营活动中,对其商品或者服务需要取得商标专用权的,应当向商标局申请商标注册”(第四条第一款)。因此,OpenAI 想要维权也有一定难度。

是否虚假服务,如何辨别?

除了使用 ChatGPT 的名称、商标外。一些公众号、小程序还会声称自己将 ChatGPT 的服务通过技术手段“转接”至国内,以方便用户使用并收取费用。那么,它们是否真的是“转接”而来的 ChatGPT 服务呢?

记者向多个国内 ChatGPT 服务提供者直接提问问题并与官方 ChatGPT 对比后发现,从回复质量上看,一些国内服务提供者确实大概率提供了面向 ChatGPT 的接口,而另一些服务提供者或许并未链接到 ChatGPT 服务,而是使用了其他 AI 对话模型。

对此,有网络安全从业者对记者表示,国内目前的“ChatGPT 服务提供方”只有两种模式:一种是取得 ChatGPT 官方接口,再提供给国内用户的“二倒手”,另一种则是完全的“山寨”,使用的 AI 对话模型质量不及 ChatGPT,自然无法给出高质量的回答。

综合北京日报、新京报、澎湃新闻等媒体报道