

上海人工智能实验室发布新一代大语言模型,更高质量语料更高信息密度筑牢基础

书生·浦语2.0会推理能共情,训练效率大幅提升

■本报记者 沈淑莎

昨天,上海人工智能实验室与商汤科技联合香港中文大学、复旦大学发布了新一代大语言模型书生·浦语2.0(InternLM2)。这个大模型包含70亿和200亿两种参数规格以及基座、对话等版本,向全社会开源,并由上海人工智能实验室提供免费商用授权。与第一代大模型相比,书生·浦语2.0在推理、代码、对话、创作等方面都有长足进步,综合性能达到同量级开源模型的领先水平。

为推动大模型在各行业的应用落地,在市经信委、市科委、徐汇区政府共同指导下,书生·浦源大模型挑战赛同时启动。首期赛事包含行业应用和创新创

意两个赛道,即日起面向全球进行场景和赛队征集。

升级数据清洗过滤技术,模型训练效率大幅提升

上海人工智能实验室主任助理、领军科学家乔宇介绍,书生·浦语2.0此次升级的核心在于回归语言建模本质,即通过更高质量的语料和更高的信息密度,筑牢大模型能力基础。

如何获得更高质量的语料?联合团队升级了数据清洗过滤技术,至今已迭代至第三代。新一代数据清洗过滤技术可实现多维度数据价值评估、高质量语料驱动的数据富集和有针对性的数据补

齐,在该技术的支撑下,仅使用约60%的训练数据就能达到使用第二代数据训练1T token(标记)的性能表现,模型训练效率大幅提升。

高质量和具有强可塑性的模型基座,是模型进行深度领域匹配的高质量起点。上海人工智能实验室领军科学家林达华说:“我们花费四个月时间提升语言建模能力,只用两周时间,基于大模型的各项应用性能就大幅提升,由此可见,语言建模能力才是大模型的硬实力。”

一次提炼30万字关键信息,“大海捞针”业内领先

修炼完“内功”的书生·浦语2.0与

多个同量级模型一起接受了综合评测,结果显示书生·浦语2.0的轻量级(70亿参数)和中量级(200亿参数)版本性能处于领先水平;200亿参数版本虽是中等规模,但在整体表现上达到了与ChatGPT比肩的水平。

在各项能力中,书生·浦语2.0的“大海捞针”本领堪称一绝。“大海捞针”是测试大模型长语境输入及理解能力的一种方法,即将关键信息随机插入几十万字的文本不同位置并设置问题,测试模型能否从中提取出关键信息。

为测试书生·浦语2.0在长文本处理任务中的能力,研究人员将一份时长3小时的会议录音转录稿输入模型,要求它从中提取出关键信息。测试结果

显示,虽然在未校对的文本中有较多错别字,但书生·浦语2.0仍从中准确提炼出了关键信息,并总结了发言人的主要观点。

林达华介绍,书生·浦语2.0能支持200K token上下文,一次性接收并处理约30万汉字(500至600页文档)的输入内容,从中准确提取关键信息,这一能力处于行业领先水平。

用户体验全面升级,会推理能共情还能创作剧本

相比于初代,书生·浦语2.0在推理、数学、代码等方面的能力提升尤为显著,综合能力领先于同量级开源模型。

“感觉每天在努力和摆烂之间摇摆,有些精神内耗怎么办?”用户输入这个问题后,书生·浦语2.0如同一位经验丰富的心理医生,给出了富有文关怀的回答:“我们每个人都经历过这样的起伏,这是正常的。当你感到精神内耗时,可以试着接纳这种情绪,就像接纳河流中的涟漪一样。不要对自己过于苛责,因为每个人都有低谷的时候。”它还能根据《流浪地球》第一部和第二部的剧情梗概,展开合理想象,创作《流浪地球3》的剧本,表现出人类面对危机时的勇气和团结。

书生·浦语2.0的对话和创作体验为何有长足进步?上海人工智能实验室联合团队解释说,这一方面得益于大模型基础语言能力的显著增强,另一方面也得益于微调技术的提升。研究人员在微调过程中,对奖励模型和对话模型进行了三轮迭代更新,每一轮均针对前一轮模型的表现更新偏好数据和提示词,不仅提高了对话的安全性,也提升了用户体验。

书生·浦语2.0的数学能力也有提升。在不依靠计算器等外部工具的情况下,在100以内的简单数学运算上能够做到接近100%的准确率,在1000以内达到80%左右的运算准确率。

全球现役装载量最大清洁能源滚装船搭载自主品牌新车驶向欧洲

7600车位“上汽安吉申诚号”定名首航



7600车位“上汽安吉申诚号”昨天开启首航。

本报记者 张伊辰摄

本报讯(记者周渊)昨天,上汽集团委托中国船舶集团建造的首艘远洋汽车运输船(滚装船)定名“上汽安吉申诚号”,并开启首航。作为全球现役装载量最大、国产化程度最高的清洁能源滚装船,上汽安吉申诚号首批搭载上汽、东风和宇通的自主品牌新车驶向欧洲。

记者了解到,该船计划常驻欧洲航线和澳新航线,年运能约3万台,将成为上汽安吉物流现有远洋运力的重要补充。这艘200米长、38米宽、13层楼高的远洋巨轮拥有7600车位,排水量超4

万吨,并创新采用LNG双燃料清洁能源,能减少30%的二氧化碳排放。以往返一次欧洲为例,预估碳排放降低1300吨,相当于50万棵树。此外,该船还拥有智能集成平台,实现全船网络及智能船舶系统的软硬件资源集中管理。

目前,上汽旗下安吉物流已建成中国最大的汽车企业自营船队,拥有各类汽车船31艘,并开通东南亚、墨西哥、南美西、欧洲等7条国际自营航线。今后3年内,7000、7600、7800、9000吨级共14艘远洋运输船将陆续加入上汽安吉物流远洋船队,积极支

持中国自主品牌加速跨海出海。

上汽集团相关负责人介绍,上汽远洋船队全力扩容的背后,源于公司海外业务快速增长。2023年,上汽海外销量达120.8万辆,同比增长18.8%。在中国汽车企业中连续8年领跑,助力我国整车出口赶超日本跃居世界第一。其中,自主品牌销量占比近92%,新能源车销量占比近24%,上汽首款全球车MG4 EV成为欧洲紧凑型纯电动车销量冠军。作为中国汽车“出海先锋”,2023年,上汽MG品牌全球销量超过84万辆,连续第五年蝉联中国单一品牌海外

销量冠军,获30余个海外权威奖项,在澳大利亚、新西兰、墨西哥、印度等多个国家跻身单一品牌前十。2024年是MG诞生100周年,品牌将全力冲刺全球年销百万辆目标。

安吉物流是上汽旗下“产业链隐形冠军”之一,经过长期发展,已成为全球最大的汽车物流公司,率先打造公、铁、水多式联运模式,业务覆盖国内近600个城市和海外100多个国家,年发运量可达千万辆规模。除了上汽旗下品牌外,安吉物流也为东风、宇通、长城等众多中国车企提供远洋整车物流服务。

冷空气蓄势待发,申城新年首雪提上日程

本报讯(记者张天弛)冷空气蓄势待发,一场雨雪降温即将抵沪。根据上海中心气象台发布的数据,在昨天暖气团团刷“存在感”、本市多区最高气温接近20℃后,一场携风带雨的强冷空气将席卷申城。据预测,本市今明两天有中到大雨,后天将迎来大幅降温,48小时

最低气温直降9℃;周末或将出现今年首场降雪。

具体来看,本轮降温稍有缓冲,分两个台阶下降。今天受降水影响,本市降温明显,最高气温仅10℃。本周末至下月初降温幅度更大,48小时最低气温降幅可达7℃至9℃;极端最低气温出现在下周二和

下周三的早晨,市区-2℃左右,有薄冰,郊区-6℃到-4℃,有严重冰冻或冰冻。

降水方面,今明两天上海有中到大雨,累积雨量15至30毫米;下周一可能出现小雪或小雨夹雪,局部地区或有道路结冰。

大风方面,周六上午起西北风开始增

大,陆地最大阵风6级,沿江沿海地区和长江口区7至8级,洋山港区和上海市沿海海面8至9级,下周三上午起风力减弱。

气象专家提醒,与前几轮冷空气相比,本轮降水过程影响范围大、持续时间长,雨雪天气对本市及跨省交通有一定影响,外出需注意安全。

(上接第一版)黄浦区人武部与中共一大纪念馆签订党建共建协议,通过基层组织共建、党员教育共育、资源互惠共享、精神文明共创、品牌文化共融等机制,共同“学理论、弘思想、强担当、铸忠诚”。松江区充分发挥老兵宣讲团和女子民兵宣讲队作用,打造宣讲品牌,在全区范围内开展宣讲活动,让爱国主义、党史国史教育、国防教育深入人心。

“党政军民学劲往一处使,拧成一股绳,上海军地在‘三共三固’活动中,锻造

出信仰信念钢铁志,绘就成党建联建工作,画出了军民团结同心圆。”上海警备区政治工作局领导介绍,真正实现让理想信念明灯点亮千家万户,让每个基层党组织成为团结群众的核心、教育党员的学校、攻坚克难的堡垒。

双拥共建深化鱼水情

为深入开展主题教育,杨浦区委、区政府协调人武部遴选推荐适合的辅导授

“三共三固”谱写军民鱼水情新篇

课对象,通过区域协作机制,一份份由辖区休干机构填报的信息表很快便反馈到人武部。

医学专业、哲学专业、马列专业……翻看厚厚一沓信息表,它们详细介绍了老干部离退休前从事岗位、身体状况和擅长的专业领域,并对老干部从事革命传统教育、专业学术讲座、医疗巡诊义诊

等活动意愿和身体状况作了预先掌握,成为地方党委、政府开展主题教育的红色资源宝库。

情系桑榆,情暖夕阳,在上海警备区所属各休干所,“三共三固”活动的炉火同样正旺。徐汇区军队离休退休干部休养所中国商用飞机有限责任公司上海某院开展教育互学,讲述老干部红色往

事,学习“商飞人”创新精神;上海警备区虹口第六离干所休干所同凉城新村街道设立全市首个“双拥会客厅”,共同编排情景党课,定期开展走访互学;上海警备区杨浦第五离干所休干所同某律师事务所达成协作,签订法律援助共建协议,共同组织普法宣讲活动。

“坚持区域性谋划,全方位融合、项目式推进,军地双方共育培塑尚武风、双拥共建深化鱼水情、双向支撑提升融合力。”上海警备区领导介绍,通过“三共三

固”活动的开展,密切了军地关系,在全市形成了资源共享、优势互补、双向促进、共同发展的军地“双向奔赴”良好氛围,推动主题教育落地生根、开花结果。

制度建设最能管根本、管长远,也是主题教育和“三共三固”活动的重要内容。下一步,上海警备区将一环接着一环抓、一锤接着一锤敲,不断固化好的经验做法,形成长效工作机制,推开面上运用、落地产生实际效果,实现工作流程规范化、创新做法制度化、成熟经验机制化。

(上接第一版)而在剧外,“或许也正是在当下的新时代,《繁花》原著才能获得茅盾文学奖,入选‘新中国70部长篇小说’,主创才有信心在电视剧领域大胆突破、尽情挥洒,腾讯出品方才有决心与魄力为这样一个‘非常规’的项目投入高成本,观众也才能看到一部前所未见的作品。”

时代让国产剧创作者敢于一部作品倾注“非常规”的投入。而对一个剧组方方面面的创作者来说,一次非典型的创作经历,或许正在重塑他们的创作观。

作家金宇澄至今记得与导演王家卫的见面之初,他陪导演看上海。在类似田野调查的过程中,作家亲眼目睹对广场舞人群、公园里孩子、一栋栋石库门建筑的高度在意。“他拿着100张石库门照

片让我挑10张最有特色的,我说看了多眼花。但王家卫导演不怕烦,他能精准辨别出每一张的细微区别。”那些日子里,他了解王家卫的创作思路与理念,包括他对物品细节的坚守、对真实还原的极致追求,以及对“回头看”的迷恋。金宇澄借用威廉·福克纳的一句话来表达王家卫与自己的相似:“人是背向过去快速奔驰的,看不清未来……过去是唯一清晰稳定的可见部分。”“保存过去的人生,保存过去的色彩。在这一块他和我的文学理念是非常相合的。”

许多主创对王家卫独有的“沉浸式”创作记忆深刻。编剧秦雯和导演阿宝见面,王家卫发问:“你觉得李李是站在至真园几楼的什么位置看着阿宝?”秦雯犯了难,毕竟连情节都还没编,角色

时代的现实感戏剧感交相辉映《繁花》似锦

站位那是下一步才会考虑的事,“但后来我知道,他是用三维的方式让我进入到画面中。”

演员胡歌也有同感,王家卫是个善于启发演员调用自己生命体验的导演。故事里,阿宝和汪小姐奔跑在南京路上,小汪的包掉落、俯身捡拾之际,阿宝驻足回望。那一刻,夕阳照射过来,布光时,导演特意关照胡歌,你望过去的小汪可能并不是小汪,也不是雪芝,“你看到的是回不去的那一段记忆”。那场戏,胡歌形容自己仿佛回到童年,看着妈妈牵着一个小男孩……当导演调用演员现实生活中最深的情感,替换角色的情绪,“刹

那间我眼泪就绷不住了”。

国产剧审美的焕新,彰显了上海的气质

《繁花》是有鲜明上海属性的。华东师范大学教授毛尖从《繁花》里视觉的上海与听觉的上海叠加,感受到了时代的现实感和戏剧感交相辉映。以汪小姐为例,初看前两集,难免会觉得她过于“叽叽喳喳”。但随着剧情推进,看到从泥潭中走出来的汪小姐,再次用她特别明媚的声音说话走路,毛尖觉得有种幸福感油然而生。“这个女孩就像上海话一样,

最新是一种声音,但在故事的推进中,她的声音变成了节奏和性格,快节奏高分贝话语变成马不停蹄的生活速率,一种沪语的轻语。这种抒情摇滚感,既升级了上海话,也升级了南方方言。”

《繁花》更是国产剧中的花开新枝丫。中国文艺评论家协会副主席、清华大学教授尹鸿将《繁花》与《漫长的季节》并置,认为这两部作品对中国的剧集,都实现了审美上的提升。“它们都超越了大众电视剧仅仅是讲好一个故事的阶段,开始追求更高的审美境界,有更丰富的审美意蕴,所以才能被广泛讨论。”

正因为有着审美焕新,中国文艺评论家协会副主席王一川和北京大学电视研究中心副主任吕帆不约而同用古典诗文来描述他们心中的《繁花》。吕帆为剧

中含蓄的情感称道,“回归了《牡丹亭》中的古典深情,情不知所起,一往情深”。王一川则从剧中看见了“流溯之美”,“它透过电影和电视剧,写实和写意、时代和跨时代、正剧和喜剧、个性和象征、进取和乡愁等方面的间性组合和交融,合力形塑出一种聚焦于上海的中国式现代化的文化形态之美,即‘流溯之美’。”

在评论家看来,那是与《诗经》一脉相承的韵味,“溯游从之,宛在水中央”,一边想向下游走,一边又想向上游回望,这种流溯之美彰显了上海的气质。“上海是现代化和现代性在中国率先发端并向全国辐射的城市,更是中国式现代化的‘领头雁’。她可以凝聚世界现代化视野中的中国式现代化景观体验,以及对于中国式现代化中新时代文化的新品位。”