



视觉中国供图

美国麻省理工学院媒体实验室研究项目显示,人工智能识别浅色皮肤男性的平均错误率不超过1%,识别深色皮肤女性的平均错误率达35%……

谁才是让AI产生偏见的幕后推手?

本报记者 谢开飞
通讯员 许晓凤 王忆希

人脸识别所导致的偏见问题一直受到广泛关注。近期,一篇关于图像超分辨率的论文引发了对于产生偏见原因的新争论。

网友利用论文中的开源代码进行了模型推

理,在使用时发现,非白人面孔的高糊照片被还原成了一张白人面孔。对此,2018年图灵奖得主、AI领军人物杨立昆(Yann LeCun)在推特上指出,当数据有偏见时,机器学习系统就变得有偏见。而这一观点,掀起一场争论浪潮。有人说,机器学习系统产生偏见不应只归咎于数据集,这种甩锅的做法是不公平的,还有技术、人类观念等原因。

识别算法中,跨种族差异确实存在广泛的准确性偏差。

柯道介绍,在算法设计的时候,往往会选择L1损失函数或者L2损失函数。若在将一个低分辨率人脸图像还原成高分辨率人脸图像的方法中,使用L2损失函数,得到的大多数人脸图像更像白人;而使用L1损失函数,得到大多数人脸图像更像黑人。这说明由于损失函数的选择不同,某个群体(黑人或白人)在一定程度上就会受算法的忽视,进而产生偏见。

这当中还存在另一种情况,假如数据带有少量的偏见,算法系统会将其放大,并变得更具有偏见。一项研究表明,如果初始数据中,下厨与女性的联系概率是66%,将这些数据喂给人工智能后,其预测下厨与女性联系起来的概率会放大到84%。一般来说,机器学习系统会选择最符合通用情况的策略,然而

算法纠偏工具已经上岗

近期,国际多家机构相继宣布禁用人脸识别技术。有色人种被还原成白人的事件,再次掀起了人们对于人工智能系统产生偏见的批判浪潮,也再次提醒人们,技术带有偏见所带来的潜在危机。

偏见可能导致无辜者蒙冤。在2018年夏天,英国媒体就报道过,由于人脸识别技术的误判,导致一名年轻黑人男性被误认为嫌疑犯,在公众场合被警察搜身。

也有报告显示,一些国外执法机构可以根据AI识别出的发色、肤色和面部特征对视频中的人物进行检索。这种技术上的偏见也在一定程度上放大了人们的偏见。

除了对人种与性别的识别偏差之外,人脸识别技术还因其在侵犯隐私方面的问题而引起担忧。2020年初,美国一家人脸识别技术公司便受到了严密的审查,有关部门发现,其面部识别工具由超过30亿张图像构建而成,然而这些图像大部分都是通过抓取社交媒体网站收集的,并且已被许多企业广泛使用。

偏见存在的同时,一批新兴的算法纠偏工

具也涌入了人工智能。

早在2018年,脸书曾发布一个算法模型,表示会在算法因种族、性别、年龄等因素做出不公正判断时发出警告以提醒开发者;2019年麻省理工学院就提出了一种纠偏算法,可以通过重新采样来自动消除数据偏见。

那么克服偏见还需要人工智能技术做出哪些努力?“开源方法和开源技术都有着极大的潜力改变算法偏见。”叶亮指出,开源方法本身十分适合消除偏见程序,开源社区已经证明它能够开发出强健的、经得住严格测试的机器学习工具。如果线上社团、组织和院校能够接受这些开源特质,那么由开源社区进行消除算法偏见的机器设计将会顺利很多。

柯道认为,可以看出偏见基本来源于数据与算法,加强对人工智能所使用的数据、算法以及模型的评价,能够在一定程度上缓解潜在的、可致偏见与歧视的因素。同时,人工智能的偏见其实都是人类偏见的反映与放大,因此人们放下偏见才是最根本的解决方案。

从居家到外出,智慧能源在惠民溪智慧能源小镇正式落成。

建筑利用了屋顶、路面等可用空间铺设光伏板,配合储能设备构建起微电网,实现建筑用能自给自足和余电上网。建筑内安装了国内企业自主研发的家庭能源路由器,根据需要自动启动热水器、洗衣机等家用电器,实现家庭能源智能管理,还安装了正在推广的新一代智能电表,居民可以通过手机随时查看每个电器的实时能耗和用电分析,参考节能建议,实现了居民用能的管家式服务。

从居家到外出,智慧能源为美好生活带来了更多惊喜。在中新天津生态城甘露溪公园,光伏座椅、智慧路灯、光伏垃圾箱、光伏路面为市民带来了全新体验。在光伏座椅的接口上为手机充电,连上智慧路灯的无线热点畅享网络咨询,这些跨界服务让市民出行不再有后顾之忧。据介绍,清洁能源发出的电量可以完全满足周边公用设施的用电需求,智慧路灯搭载了视频监控、PM_{2.5}监测等多项功能,为服务社会治理和生态民生提供了新的经验。

而在中新友好图书馆公交站,充电机器人准确识别无人公交车停靠,自动完成插枪、充电、拔枪全过程。

第四代人工智能配网带电作业机器人也于近日正式在滨海新区上岗作业了。它运用了三维环境重建、视觉识别、运动控制等核心技术,实现了机器人自主识别引线位置、抓取引线、完成剥线、穿线和搭火等工作。

现在还没有100%无偏见的数据集

随着人工智能时代的到来,人脸识别技术正被越来越广泛地应用在执法、广告及其他领域。然而,这项技术在识别性别、肤色等过程中,存在着一定偏见。

在美国,人脸识别系统中白人男性的识别率最高,而黑皮肤女性的识别率则最低;在亚洲国家的人脸识别系统中黄种人的识别率相比白人种人的识别率差距就会小一些。“福州大学数学与计算机科学学院、福建省新媒体行业技术开发基地副主任柯道博士说。

这次陷入争论中心的人脸识别偏见问题其实是一个存在很久、热度很高的话题。

早前,美国麻省理工学院媒体实验室主导的一项名为《Gender Shades》的研究项目在研究了各个前沿科技公司的面部识别系统后发现,各系统在识别男性面孔与浅色面孔方面表现更佳,浅色皮肤男性的平均识别错误率不超过1%,浅色皮肤女性的平均识别错误率约为7%,深色皮肤男性的平均识别错误率约为12%,深色皮肤女性的平均识别错误率则高达35%。

那么本次争议中,杨立昆指出的数据集问题是如何产生偏见的?有无100%没有偏见的数据集?“不均衡的数据,导致了系统在学习时,更好地拟合了比重较大的数据,而对占比小的数据的拟合则表现不佳,这就可能最终导致了偏见的产生。”柯道介绍说,如一个常见的人脸识别数据集LFW,数据集中有近77%的男性,同时超过80%是白人。

“目前,几乎可以肯定的是没有100%无偏见的数据集。”柯道举例说,在收集数据时,可以针对性地让男性与女性数目一样,甚至也可以让各种肤色的数据数目都一样。但是,其中可能大部分人是没有戴眼镜的,这可能就对戴眼镜的人进行识别时带来了偏见。进一步讲,各个年龄段的人脸也很难做到完全一致,导致最后学习到的模型可能对戴眼镜、年长或年幼的人脸产生偏见。所以有时候讨论数据偏见的产生,会发现其产生的原因是考虑问题的角度不同。

出现偏见不能只归咎于数据

众多专家也指出,数据可能是机器学习系统中偏见的来源,但不是唯一的来源。

其实,在本次争议中,杨立昆在推特中已经很明确地表示,他的本意并不是说问题都出在数据集上,只是针对这篇论文,其算法在特定场景下,换个数据集就能解决问题,但并不是说所有的机器学习系统偏见都来源于数据集。

那么,谁才是产生偏见真正的“幕后推手”?

AI学习过程应该是架构、数据、训练算法等的相互作用。算法包括了特征的选择、损失函数的选择、训练方法的选择以及超参数调节等,其本身的缺陷也是偏见产生的一个原因。

2019年12月,美国国家标准技术研究院进行的一项研究发现,在大多数当前评估的人脸

挖掘能源数据,让智慧城市“心中有数”

本报记者 陈曦

“能思考、会分析”的能源大数据平台,能够提供互联网无线信号的智慧路灯,可以进行电能消耗分析的智能电表,把电动汽车变为“电网充电宝”的充电桩,替代人工作业的配网带电作业机器人……在第四届世界智能大会上,这些陈列在“云端”的电力“黑科技”,展示着能源互联网的前沿技术。能源领域智能创新,正有力服务着经济社会发展,改变着我们的生活。

能源大数据为城市管理赋能

近日,在天津市能源大数据中心的主屏幕上,“电力数据看经济”的电量曲线逆势上扬,5月份

天津投产项目的电力容量同比增长89.68%,新增项目的申请电力容量同比增长202.41%……为用好电力数据这张经济发展的“晴雨表”,国网天津电力搭建“能源大数据运营管理平台”,对能源数据进行精加工,让智慧城市“耳聪目明”。

目前该平台已汇聚了智慧能源小镇、车联网平台、园区、企业等多方面的数据,应用大数据、云计算、物联网等信息技术和智能技术,打造面向政府的能源决策支持平台、面向企业的智慧能源服务平台,满足多样化需求。

如今,天津市能源大数据中心已经推出服务政府部门、服务用能企业、服务居民用户、服务能源系统4大类共30项数据应用成果,提供定制化服务。其中“电力数据看经济”“一区一行一指数”

“应急减排企业监测”“住房空置率分析”“企业信用评价”“家庭能效分析”已经成为服务于政府、企业和市民的名星产品。

助力节约楼宇用能成本

整齐排列的光伏板、迎风而立的风机,让天津北辰区大张庄智慧能源小镇充满了现代气息,阳光、微风、地热……大自然的赠予源源不断地转化为生产生活所需能源,奔流涌动。

“我们搭建了智慧能源管控平台,让客户以最低的用能成本进行生产办公。”国网天津城东公司客户经理杨国朝介绍,在智慧能源小镇的商务中心,近40%的能源均由可再生能源自产自用,配备的储能设备不仅可以解决分布式电源出力不稳定性问题,而且可以利用夜间低谷电价蓄能,在白天用电高峰时供电,通过峰谷电价差为客户节省一笔电费支出。同时,智慧能源管控平台实时监测大楼房间的温度变化,自动调节地源热泵的冷热出力,每年可以为楼宇节省用能成本100多万元。

在优化能源供给与消费的同时,小镇内打造了国内首个智慧能源工厂,根据生产线的设备特性、电压敏感度、影响损失等因素,实时分级治理电能质量,对精密加工生产线加装智能控制系统,可以实现毫秒级消除电压波动,降低了废品率,企业每年可以增收100多万元。

实现用能的管家式服务

日前,在中新天津生态城,一座充满未来气

息、科技含量十足的零能耗智慧建筑在惠民溪智慧能源小镇正式落成。

建筑利用了屋顶、路面等可用空间铺设光伏板,配合储能设备构建起微电网,实现建筑用能自给自足和余电上网。建筑内安装了国内企业自主研发的家庭能源路由器,根据需要自动启动热水器、洗衣机等家用电器,实现家庭能源智能管理,还安装了正在推广的新一代智能电表,居民可以通过手机随时查看每个电器的实时能耗和用电分析,参考节能建议,实现了居民用能的管家式服务。

从居家到外出,智慧能源为美好生活带来了更多惊喜。在中新天津生态城甘露溪公园,光伏座椅、智慧路灯、光伏垃圾箱、光伏路面为市民带来了全新体验。在光伏座椅的接口上为手机充电,连上智慧路灯的无线热点畅享网络咨询,这些跨界服务让市民出行不再有后顾之忧。据介绍,清洁能源发出的电量可以完全满足周边公用设施的用电需求,智慧路灯搭载了视频监控、PM_{2.5}监测等多项功能,为服务社会治理和生态民生提供了新的经验。

而在中新友好图书馆公交站,充电机器人准确识别无人公交车停靠,自动完成插枪、充电、拔枪全过程。

第四代人工智能配网带电作业机器人也于近日正式在滨海新区上岗作业了。它运用了三维环境重建、视觉识别、运动控制等核心技术,实现了机器人自主识别引线位置、抓取引线、完成剥线、穿线和搭火等工作。

瞭望站

张文宏:疫情防控中,AI的真正作用在于物资调配

侯树文 本报记者 王春

近日,2020世界人工智能大会云端峰会拉开帷幕。大会邀请复旦大学附属华山医院主任张文宏作为对话嘉宾介绍人工智能在应对疫情当中的作用。

“疫情中人工智能发挥了什么作用?未来公共建设中,人工智能还能发挥什么作用?人工智能可以帮你预测疫情进展吗?”清华大学国家金融研究院院长朱民抛给张文宏的几个问题,同样是公众关切的问题。

当社区封闭以后,人工智能在智能化管理中起到了重要作用。“社区都关闭以后,整个城市怎么运行,市民怎么生活,又怎么维持一到两个月时间?这次,人工智能帮助政府做到了智能化管理。”张文宏为AI点赞。

张文宏说,作为医生,他很不愿意被大数据替代掉,也不愿意拥有大数据后,有人把X光片数据一输入就能开出处方。针对此次疫情,他认为,人工智能技术真正发挥作用的地方不是在医疗诊断治疗方面,而是公共卫生体系的资源调度和智能化运营,“从人工智能参与到物资的调配的时候,我们就意识到将来人工智能具有非常强大的物资调配能力”。

他特别提到,要反思疫情防控公共卫生存在数据“孤岛”问题:医疗机构的数据如何第一时间汇聚到大数据平台?线上数据和线下数据如何融合,然后预警,构成一个完整的体系?“今天飞来多少个航班,航班所在国家的疫情如何,我们的医疗资源要相应增加到什么程度,这些都应该有真实的数据。今天我们的防控是熔断式的,熔断就不能‘人工智能’。将来的数据应该是真实世界的真实数据,非常期待线上线下的数据结合。”

“大家现在对人工智能寄予了厚望,但是从新冠肺炎疫情开始至今,我感觉这仅仅是它的一个起点。”张文宏表示。

情报所

复杂网络关键参与者 快速识别算法实现新突破

科技日报(程光权 凡宁 记者张强)发现复杂网络中的关键参与者,对认识及优化网络整体效能至关重要。记者近日从国防科技大学获悉,该校系统工程学院研究人员创造性地提出了一种名为FINDER的深度强化学习AI算法,实现了对复杂网络中关键参与者的快速准确识别,在效果、性能及普适性等方面均超越了现有的解决方案。相关研究成果近日在《自然·机器学习》发表。

在复杂网络中,如果节点数增加,寻找关键节点的时间会呈指数级增长,这在计算机科学中被称为NP-hard问题,是优化算法领域的终极挑战。解决这一问题的传统解法包括精确算法、近似算法、启发式算法等,但这些算法在准确性和计算效率上难以取得令人满意的平衡。更重要的是,目前缺乏这一类问题的统一求解框架,以致同一类问题的不同应用场景都需要专门设计不同的算法。

据介绍,该校研究人员此次提出的FINDER是求解这类问题的统一算法框架。它能够在经典模型生成的小型合成网络中先行离线训练,而后根据特定问题场景奖励函数的指导,自主学习掌握“聪明”的选点策略——根据当前状态(即当前观察到的网络结构),选择能够获得最大预期回报的行动(即应选择的节点)。

多个大规模真实网络上的实验结果表明,与现有技术相比,FINDER在寻找复杂网络关键参与者的准确性和计算效率上均取得了更好的表现。特别是在效率上,可以轻松扩展到百万节点级的大规模网络。此外,FINDER还是一个高度灵活且通用的框架,只需更改其奖励函数,就可以应用于不同的问题场景。这为分析复杂网络的组织结构原理提供了新的分析范式。

据悉,该算法未来有望在人群流行病控制、药物的合理设计、疾病致病基因识别、社交媒体舆论引导及谣言阻断等方面发挥重要作用。

上海张江建设“机器人谷” 推动智能机器人走进生产生活

据新华社(记者何欣荣 滕佳妮)快速发展的人工智能技术,推动机器人产业不断升级。上海张江集团最新表示,将携手合作方,在张江科学城建设“机器人谷”,推动智能机器人走进生产生活。

专注于机器人产业的赛道投资,近日与张江集团签署战略合作协议,共同组建服务机器人生态联盟,打造“AI+机器人”公共服务平台。“天上飞的、地上滑的、水上漂的,各种机器人企业都面临难得的发展机遇。以场景应用为导向,未来智能机器人和服务机器人大有可为。”赛道投资负责人石东华说。

人工智能技术与机器人产业的结合,是当前的一大看点。迈宝智能科技创始人魏巍说,运用人工智能与机器人交互技术,可以为工业物流、建筑搬运等行业的体力劳动者,提供运动辅助型的可穿戴式外骨骼产品,促进相关产业发展。

张江集团董事长袁涛介绍,目前在张江科学城,北有张江人工智能岛,南有康桥核心智造园,“北创新,南智造”的格局,为智能机器人产业发展构建了良好的生态环境。

作为上海重要的机器人产业集聚地,张江拥有多家国内外机器人企业。国际机器人巨头ABB将中国总部设在张江,国内机器人领军企业优必选以及钛米机器人、微创医疗机器人等一批企业也在张江集聚。

上海市机器人行业协会发布的报告显示,从机器人整机、核心零部件到系统集成应用,长三角的机器人产能占全国50%以上。在推动传统制造向智能制造升级方面,机器人扮演着越来越重要的角色。



在中新友好图书馆公交站,充电机器人准确识别无人公交车停靠,自动完成插枪、充电、拔枪全过程。

受访者供图