

“阿尔法风”上岗,天气预报或精确到1公里内

好机友

本报记者 张晔 通讯员 林雯

“15分钟后,公园门口的麦当劳很可能会下雨,你在手机上会收到服务提示。”未来,无论你身在城市哪个角落,都可能收到所处1公里甚至更小范围内的天气预报。

刚刚过去不久的世界气象日,“智慧气象”成为今年的主题。科技日报记者从南京信息工程大学获悉,该校信大气科学研究所团队(南京信大气科技有限公司)正在使用人工智能预测天气发展,特别是强对流天气的发生发展,这套已经在全国多地气象部门“实习”的“AlphaWind”(阿尔法风)系统,将更为准确地预

测雷暴大风等灾害性天气。

3月15日,冷空气裹挟八级大风南下,在19—20时之间,南京南部的降水1小时降雨量达66.3毫米,为罕见的强度极大的短时强降雨过程。“当天傍晚,‘阿尔法风’系统就已经测算到可能出现雷暴大风的局部区域,并向气象预报员们发出了警告。除了南京以外,华东地区特别是江苏上游的安徽省境内,出现的一系列局部雷暴大风也被它迅速捕捉,一一标出。”团队带头人苗春生教授举了个例子。

记者在现场看到了安徽某地即时的气象雷达图,当图像被输入到构建的卷积神经网络模型中后,系统快速标识出了雷暴大风可能发生的红色区域。“‘阿尔法风’目前主要针对气象预报领域里的难题——短时强对流天气,包括短时大风、强

降雨、强雷暴等灾害性天气。”苗春生解释,气象行业利用的原本就是“大数据”,过去,预报员凭借丰富的经验,结合多年的资料数据进行预测。但对于关键转折性天气,预报员很难即时反应,这也使得暴雨、龙卷风、冰雹等极端天气成了气象业务的关键瓶颈。

团队核心成员、强天气预报专家邵玲玲研究员从事气象预报工作二十几年,是复杂气象方面的专家。她告诉记者,多年前,气象预报员遇到诸如疑似雷暴这样强天气因子时,都是一群人围在一起讨论、判断,一般要30分钟才能基本准确辨别,经常会空报、漏报,准确率并不高。而现在,“阿尔法风”人工智能系统不需要针对不同地区、不同季节的极端天气统计总结出规律或特征,只要“投喂”历年海量的天气数据记录,

“教”给它各种算法模型,不断纠偏,就可以日进千里,越来越准确、迅速地对各种极端天气加以识别。

人工智能不仅保证了气象预报的时效性,未来还能准确预报1公里的小尺度超局地天气。该技术还将在高铁、海洋工程等大型项目上发挥气象保障作用。团队目前正在和中国高铁总公司进行合作研究。“在地形、受力各方面复杂影响下,如何每隔10秒、间隔数公里报一次风向风速?这是过去传统方法不能解决的。”苗春生介绍,“通过大数据深度挖掘高铁路段资料,我们的预报准确率目前可达84%。”除了高铁,智慧气象还可以嵌入自动驾驶系统。“目前有国际汽车公司希望我们能够在团雾、道路结冰等恶劣天气下,利用气象大数据提高自动驾驶的可靠性。”苗春生说。



数据挖掘大战 机器怎么做到「见信如面」

本报记者 张佳星

大数据杀熟?隐私换便捷?一度被热捧的大数据挖掘,近日站在了舆论的风口浪尖;一些商家利用大数据挖掘技术“杀熟”被网友亲测证实;百度董事长兼CEO李彦宏一句“中国人对隐私问题没有那么敏感”,更是让它的处境雪上加霜。大数据挖掘技术就像一位有了负面新闻的明星,霎时间光彩黯淡,似乎变成了偷人隐私的小贼。

用户画像:机器给人类贴标签

“通过打标签的方式建立用户画像,是数据挖掘常用的一种技术。”北京大学计算机科学技术研究所多媒体信息处理研究室主任彭宇新教授解释,建立用户画像就是利用社交网络的信息,根据用户社会属性、生活习惯和消费行为等信息,抽象出一个标签化的用户模型,目标是使机器实现类似于人的“见信如面”的能力。社交网络数据是实现这一目标的基础,机器对人的“初相见”多是源自于对社交网络数据的挖掘。

标签,通常是通过对用户信息进行分析得到的高度精炼的特征标识,使得机器方便做信息提取、聚合分析等处理。标签本身无需再做过多文本分析等处理工作,这为利用机器提取标准化信息提供了便利。

“有了标签,计算机就能够自动处理与人相关的信息,能够通过算法、模型逐步‘理解’人。”彭宇新介绍,多个标签共同完成画像,整个过程可分三步走:一是采集数据,即基于文本的

《大数据时代》一书畅销之后的几年,大数据虽不再那么当红,但并未隐退,它的持续发展已成为人工智能得以实现的基础之一。

那么,大数据挖掘究竟是怎样的技术?从诞生发展至今,那些埋头苦干的技术人员又让它长了哪些本领?面对大数据难以管理的问题,有没有技术手段加以控制?

信息抓取,口语称为“爬数据”;二是用户行为建模,通过机器学习技术,形成算法模型,判断用户可能的一些行为;三是可视化展现,把机器运算出来的结果,通过能让人理解的方式展现出来。这三步是多轮调整的,在实际应用中,根据结果的反馈,以及业务需求,可能进行二次建模等调整。

整个过程的影响参数是相对多元的,不同的行为类型,对于标签信息的权重影响也不同。以应用最广的商品营销为例,比如网售红酒,如果“购买”权重计为5,“浏览”计为1,加上浏览间隔、驻留时长、生活习惯等,通过复杂的算法最终呈现出一个标签的权重,再形成画像。

基于用户画像技术,大数据挖掘进行分类和关联规则计算等分析:例如喜欢红酒的用户有多少,喜欢红酒的人群中,男、女比例是多少,喜欢红酒的人通常喜欢什么运动品牌等等。

跨媒体智能识别:为计算机装上慧眼

“以前文本信息占主流,现在图像、视频等多媒体数据铺天盖地而来。”彭宇新说,后者目前占据大数据的80%以上。

数据类型发生的巨大变化,使得智能识别的任务更加艰巨。“管不住”和“用不好”的问题日益凸显。“机器只能读懂自己的语言。”彭宇新说,人类世界的所有语言都要转化为机器理解的语言才能被识别,以前只处理文本相对简单,而现在要加上复杂的图像、视频等数据。

“例如,世界上有数千种鸟类,很多种的差异非常细微,即使是有专业知识的人类也很难准确辨认,计算机自动识别的难度就更大。”彭宇新说,图像、视频内容理解的难点在于如何进行语义自动识别,这也是他们团队多年攻关的课题之一,为此团队发明了基于注意力模型和深度学习识别的方法。

注意力模型,顾名思义是让计算机自动定位图像的显著性区域,以此提高检测精度;深度学习,是指计算机能够利用已经学到的知识加速对新知识的识别,同时通过动态扩容以支持新概念的检测。

新模型新算法的发力,帮助机器快速识别图像、视频的语义信息。彭宇新团队近年来六次参加国际权威评测TRECVID的视频样例搜索比赛均获第一名,并在与卡内基梅隆大学、牛

津大学、IBM Watson研究中心等参赛队伍的较量中胜出。其中一个题目就是在464个小时的视频中快速准确地找出所有的伦敦地铁标志,彭宇新团队仅用了不到1秒就成功胜出,获得第一名。

单媒体信息的分析与识别之上,如何进一步让机器像人类一样能看、能理解呢?

为达到跨媒体信息融合与一体化分析识别的目的,项目团队首先把数据按照不同媒体类型自动分发到对应的分析与识别模块。例如,对视频镜头进行分割,对关键帧进行提取,然后分发到镜头检索、片断检索、视频字幕识别等模块中,对单媒体分析结果进行跨媒体语义关联分析,实现跨媒体信息的语义协同。“一种常用的方法是构建第三方空间进行跨媒体关联。”彭宇新说,“计算机根据我们教它的模型分别为图像、视频、文本、音频抽取表征,再共同投射到一个第三方空间中,这样不同媒体的信息就可以对话了。”

技术的“抽丝剥茧”,让图像、视频中的信息可以如文本一般清晰透明。“我们是瞄着应用去的,准确率、处理速度都经过多年的优化,已经可以进行实际应用了。”彭宇新介绍,这项技术不仅帮助新闻媒体等行业进行数据管理和检索,还在助力互联网管理部门对大数据进行分析与监测。

新鲜事

苹果“偷走”谷歌AI总管 硅谷上演重量级人才战



据《纽约时报》报道,近日,谷歌原来的人工智能及搜索负责人约翰·詹南德雷亚,将加盟苹果,领导其机器学习和人工智能的发展。

现年53岁的詹南德雷亚负责谷歌人工智能产品方面的工作,包括对互联网搜索、Gmail和谷歌助手等进行整合。他的加入对于苹果公司来说是一个巨大的收获。多年来,苹果一直致力于人工智能领域开发,比如在计算机视觉和自然语言处理领域。据《泰晤士报》报道,苹果公司于周二发表声明,詹南德雷亚将负责苹果的“机器学习和人工智能战略”,并成为直接向苹果首席执行官库克汇报工作的16名高管之一。

随着詹南德雷亚的加入,苹果或许能够招募到更多的顶级人才,并改进其算法。而对谷歌来说,失去詹南德雷亚可能不会对其人工智能的发展产生重大影响。接替詹南德雷亚的是被业界广泛认为极有才华和极受信任的杰夫·迪安。

韩国拟研发“杀人机器人” 周志华等50名学者联名抵制



据网易智能报道,近日,韩国科学技术院因为与军工企业“韩华Systems”合作研发自主武器,遭到来自全球20多个国家的50多名AI学者联名抵制,其中包括Yoshua Bengio、Geoffrey Hinton以及南大计算机系教授周志华。

这50多名学者通过新南威尔士大学官网发出一封公开信。他们担心,韩国科学技术院研发自主武器将会加速AI军备竞赛,像这样的技术可能导致更大规模的战争加速到来。他们在信中警告,如果成功研制自主武器,战争速度之快、战争规模之大,将是前所未有的,也有可能成为恐怖分子的武器。

学者们在公开信中表示,他们将抵制与韩国科学技术院的任何合作,也不会接待该院派出的访客,并停止所有与该院的接触,直到该院院长承诺不会研发无人操纵的自主武器。

公开信发出后,韩国科学技术院总裁Sung-Chul Shin表示,听到抵制的消息他感到很难过。“我想重申,韩国科学技术院没有任何意图参与开发致命的自主武器系统和杀人机器人。”他在一份声明中说。

港口实现无人化 图森将推智慧港口方案



据网易智能报道,近日,L4级自动驾驶货运卡车企业图森未来对外发布基于完全自有车辆的港区内部集装箱卡车车队自动驾驶测试视频,宣布进入港内集装箱卡车无人驾驶运输市场。

图森未来表示,经过在北方某深水集装箱港口内的无人内集卡车队保密测试,他们已实现对接现有港机系统及港务系统的无人码头内集装箱转运车队解决方案,拥有了由干线运输到区域内运输的整套无人驾驶技术解决方案,目前图森有5辆自有自动驾驶卡车在某港口测试,预计年底达到25辆。

此外,图森未来透露,目前已有多家港口向他们表达了合作意愿,希望能积极合作,打造基于人工智能和自动驾驶技术的智慧港口。

(本版图片除标注外来源于网络)

探秘 AI 纺织车间:可在黑暗中“穿针”

产业界

本报记者 王延斌

借助AI技术,原本专业化的知识可为普通人掌握,原本需十人完成的工作现在只需要一人。在世界最大纺织企业、“世界500强”魏桥创业集团的车间里,随处可见的AI技术不炫技玩虚的,而是瞄准问题,跟生产过程紧密联系,源源不断地产生价值。

把棉花纺成纱,一般要经过清花、梳棉、条并卷、精梳、并条、细纱、包装等主要工序。过去,这些工序依靠“人海战术”,动辄数百名纺织工围着一堆机器转。但高科技的介入改变了传统模式——得益于12.5万纱锭智能化生产线,这里

所有的纺纱流程都实现了自动化。

进入车间,排列整齐的各色设备运转有序,却少有人走动,取而代之的是随处可见的“机器人”,它们正成为车间里的主角。例如最直观的包装环节,作为纺纱工业的主要环节之一,包装原本是“苦活、累活”的代名词。但这个车间建立起长度为35公里的轨道,粗纱从一楼到二楼自动入库,整个流程无任何人工直接参与。

看得见的改变背后,是看不见的大脑在起作用。

“在整个生产线上,我们安装了15万个传感器,对所有的一线数据进行采集,然后进行智能分析,其结果可反馈给系统自动调节生产过程。”魏桥纺织董事长张红霞向记者举例:像细纱车间工人,现在一个班只有6个人,一

个人要看管2万纱锭,这是以前想都不敢想的,因为以前需要60人,一人最多看管2—3千纱锭。一个更直观的数据比较是:纺织全行业的每1万锭所用工人数量为100人,但在魏桥纺织最少只用10人。

装上了“智慧大脑”的智能纺纱生产线到底有多聪明?魏桥纺织生产管控中心主任陈以军告诉记者:“我们利用大数据、云计算、互联网等技术,对纺纱每个工序和作业点进行可视化监控,并通过互联网跟踪订单的进展和质量情况,同时,将生成的实时数据进行集成和分析,以此反作用于生产管理。”

在纺纱行业存在着一种现象:纺纱对车间环境要求高,比如需要灯光明亮,车间温度时刻保持28摄氏度以下,湿度保持55%—65%以内,以保证

产品质量。“生产全程自动化”“控制系统智能化”“在线监测智慧化”,这“三化”的实现解决了上述两大难题,甚至可实现“熄灯生产,织造不断头”的生产状态。如同在黑暗中练就的“穿针眼儿”绝活,不借助灯光的高科技,完成最熟练的纺织工也完成不了的种种动作,从而打造出全球纺织业的第一个“熄灯车间”。

产业转型升级,向AI技术要“红利”是当下很多产业的发展主流。但事物发展总是两面性,不好的一面便是,随着生产效率提高,会有大批的劳动力被精简。“纺织行业里,90%的就业人口是农民工。我们不能以智能化为借口让工人下岗。”魏桥创业集团董事长张士平曾向记者坦言,新技术推动产业升级,他们会扩大规模,吸收这部分被精简的劳动力,不会让一个纺织工人失业。

扫一扫
欢迎关注
AI瞭望站
微信公众号

