



视觉中国供图

人工智能让舆情监测 从信息检索走向内容多维度识别

◎本报记者 华凌

全民互联网时代,舆情无时不有,风险无处不在。目前,中国网民总体规模已超过10亿,网络成为人们表达个人意见和观点的重要渠道,社会舆论也随之从线下转移到线上。网络舆情信息成为舆情信息的重要组成部分,给全媒体时代做好舆情工作带来全方位的挑战。

“智能化舆情管理是行业发展的必然。伴随着自媒体、移动端的快速发展,在不到10年的时间里,每日舆情数据总量从百万级扩展到千万级,当前已经达到亿级。如此庞大的舆情信息,远远超过正常阅读的极限,要对海量的信息进行分析,势必使用人工智能技术对舆情信息进行处理和研判。”2月8日,人民在线副总编辑、人民网新媒体智库高级研究员刘鹏飞在接受科技日报记者采访时表示。

AI舆情系统优势愈发突显

与传统舆情系统相比, AI舆情系统有何不同?

人民在线副总经理汪大伟答道, AI舆情系统是指以人工智能技术进行中文全媒体资讯自动化分析与大数据情报挖掘的综合系统。与传统舆情系统相比,其有两个明显的转变:舆情监测从“检索”到“算法”的转变;数据计算从“简单计算”到“深度学习”的转变。AI舆情系统的优势在于可以精确地对文本类信息进行实体识别、语义消歧、知识图谱构建、话题分类、自动摘要、情感分析,并对于图像类的信息进行有效的品牌识别、人脸识别、物体识别和文字识别等。

汪大伟进一步解释,随着大数据、云计算、人工智能技术的发展与用户量的日益增长,舆情智能化发展亟待解决两个需求:以移动化应用为代表的“浅舆情”需求以及需要深度分析的“潜舆情”需求。而无论是“浅舆情”还是“潜舆情”都需要一个强大的舆情数据处理平台。

首先, AI舆情系统提升了数据的精准性。早期的舆情监测,通常由“关键词”搭配“与、或、非”的判断逻辑进行数据检索,往往需要辅以大量的人工,对数据进行二次处理。而智能

化的监测,则通过自然语言处理技术对内容进行多维度识别,从而提升数据的准确性。基于自然语言处理技术, AI舆情系统运用垃圾分类模型提升数据精准度,并通过情感分析技术获取敏感信息,实时表现舆论状态,评估舆论走向。在此基础上,系统还能通过事理图谱、热点聚类、文本分类等学习方法,对舆情事件的发展脉络、特征分布、风险等级进行自动阶段性总结,并给出趋势预测。

其次, AI舆情监测的优势,还体现在对“非确定”信息的监测上。长期以来,舆情监测一直存在一个难题,即对一个部门、企业、话题、事件等确定性的“主题监测”相对容易,但对一些“无主题”的监测,则困难得多。

例如,如果想关注全国的校园安全舆情并精准推送给相关部门,很明显这很难穷举全国所有学校的关键词及安全相关的关键词。而AI舆情系统通过设定主题地域属性、媒体权重、热度属性、行业属性、情感属性、敏感属性,可以快速将全国的舆情事件进行定位、定性,并在此基础上进行信息的抽取、挖掘、聚类和分,为关联单位提供智能监测、智能预警、智能研判等服务。

率,优化市场环境。

汪大伟指出, AI舆情系统提升了信息概括总结的能力,舆情作为重要的社会评价数据,在跟其他维度的数据进行融合之后,再次扩展了应用场景和使用价值。很多政府机构将舆情数据视为社会治理数据的重要组成部分;企业则常将

AI舆情系统提升了数据的精准性。早期的舆情监测,通常由“关键词”搭配“与、或、非”的判断逻辑进行数据检索,往往需要辅以大量的人工,对数据进行二次处理。而智能化的监测,则通过自然语言处理技术对内容进行多维度识别,从而提升数据的准确性。

舆情数据作为大数据风控和智慧营销的重要参考指标。显然, AI舆情系统推动了“网络问政”向“网络理政”的转变。

随着近些年数据风控应用越来越广泛,在治理个人征信数据滥用乱象之后,很多金融企业与舆情数据作为重要的风险指标,希望利用舆情数据进行风险识别、风险估测和风险评价等。

汪大伟说:“AI舆情系统不仅能够将已有风险归纳为经验知识,还可以利用人工智能技术和知识图谱功能,将某一种经验扩充为某一类经验,以此来实现对未来风险的精准预测。此外, AI舆情系统的知识图谱功能,具有推理计算的能力,可以发现以往并未存在但将来可能存在的

风险,满足用户的深层需求。”

2020年9月,《关于加快推进媒体深度融合发展的意见》指出要以先进技术引领驱动融合发展,用好5G、大数据、云计算、物联网、区块链、人工智能等信息技术成果,加强新技术在新闻传播领域的前瞻性研究和应用。

在传统媒体时代,一篇新闻稿件发布的同时,意味着工作即结束。而在融媒体时代,稿件的发布仅仅是稿件传播的开始。通过对稿件内容的分析,可以更加精准地分析出稿件更适合在什么时间发布、更容易在什么平台上传播、更适合用哪种表现形式。而这一切,很可能是基于舆情数据对当前新闻舆论的精准分析。

延伸阅读

媒体智能化初步形成三大模式

2月5日起,北京广播电视台在播出北京冬奥新闻时,屏幕左下角多了一个熟练运用手语播报的“虚拟人”——“冬奥手语播报数字人系统”,让听障人群更好地感受到了冬奥盛况。从虚拟新闻主播到手语播报虚拟人等,人工智能技术已经成为北京冬奥会新闻播报的一个重要辅助技术,而智能化也早已成为如今新闻传媒发展中的一个关键词。

中国传媒大学等联合发布的《中国智能媒体发展报告(2020—2021)》指出,目前媒体智能化已初步形成三大模式,即智慧广电模式、智慧报业模式、商业平台智能化模式。

在强有力的政策引导与扶持下,我国智慧广电得以迅猛发展,人工智能技术不断融入广播电视各业务流程,催生出了智能视频修复、智能视

频编辑、广电网络智能分发、广电云监管、多维度审核体系、广电公共应急信息智能服务等一系列智慧广电创新应用。

在人工智能、媒体融合相关政策助推下,近年来中国报业智能化转型进程显著提速,目前已进入媒体融合“深水区”。整体来看,传统报业智媒建设经历信息生产单点环节赋能、采编安全链路智能支撑、中台驱动平台级产品对外输出3次进阶。

头部互联网商业媒体平台在市场驱动下率先开启智能化升级转型,目前人工智能技术覆盖从汇聚、管理、生产到发布的全链路,智能应用呈现出百花齐放的局面。在疫情催化、全面视频化的趋势下, AI+视频也成为商业平台角力的新赛道。

从“网络问政”到“网络理政”

近年来,有不少地方的市场监管部门越来越重视舆情数据的使用。将当前舆情压力最大、最敏感的行业作为工作关注的重要方向,特别是对一些周期性、常态化的问题提前布局,通过精准识别、精准执法,以监管促合规,在降低自身舆情风险的同时,也提升工作效率

深度学习新方法 解决定量光声层析成像缺乏训练数据难题

◎本报记者 陈曦 通讯员 赵晖

日前,天津大学精密仪器与光电子工程学院李娇副教授和高峰教授团队,利用定量光声深度学习的方法,实现了活体深层组织的光学功能“真实透视”成像,这在世界尚属首次。这将为获取活体组织生理病理相关的血氧特性图像提供高空间分辨率定量成像方法,可用于肿瘤早期筛查、良恶性诊断以及抗癌药物疗效在体监测与量化评估。该论文近日发表在国际光学期刊《Optica》上。

深度学习新方法进入医学影像学

定量光声层析成像是一种新兴的无创生物医学成像技术,可结合传统光学成像的功能性以及传统超声成像的高清晰度等优势,直接获取深层组织光学吸收系数图像,因此受到国内外研究机构和医疗企业的广泛关注。

但目前的定量光声层析成像方法需要庞大的计算资源和时间消耗,且存在稳定性差、先验信息依赖性强及误差大等问题。

近年来,深度学习方法进入医学影像学领域。但深度学习想要实现既定功能一般需要两个过程:训练过程以及实际识别过程。

“如果想要让神经网络充分发挥其学习能力,必须有大量带有标注的真实数据用于其训练过程。”论文第一作者李娇表示,然而,在许多生物医学成像中很难获得深层组织特别是活体组织的真实值(如光学吸收系数),因此很难构建大量带有标注的真实实验数据集用于深度神经网络的训练,造成深度学习在许多生物医学成像领域难以应用推广。

针对上述难题,李娇和高峰团队首次提出了无需标注真实数据的定量光声层析成像深度学习新方法,实现了深层组织吸收系数的准确重建。

解决深度学习方法的泛化性难题

该研究的创新点之一是解决了神经网络训练数据问题,利用风格迁移网络(SEED-Net)实现仿真数据与实验数据的无监督自由转换,将丰富的带标注仿真数据转换到实验域,即生成了大量带标注的“实验数据”用于后续神经网络训练。

该研究的创新点之一是解决神经网络训练数据问题,利用风格迁移网络实现仿真数据与实验数据的无监督自由转换,将丰富的带标注仿真数据转换到实验域,即生成了大量带标注的“实验数据”用于后续神经网络训练。

“我们提出的SEED-Net不仅可以解决定量光声层析成像方面缺乏真实数据集的问题,在其他生物医学成像领域例如光学、荧光层析成像中同样受限于缺乏足够的带标注的真实实验数据,也可使用该网络通过丰富的仿真数据生成‘实验数据’,进一步发展适用于实际应用的生物医学成像深度学习新方法。”李娇说。

该方法具有普遍适用性,适合在不同光声成像系统、其他光学成像技术和整个生物医学成像领域推广及应用。

“这也从一定程度上解决了深度学习方法的泛化性难题。”论文的通讯作者之一、天津大学副教授孙彪介绍说。

该研究的另一创新点是结合实际光声数理模型设计了一种双通道神经网络,分别考虑组织光强分布和光学吸收系数对初始声压图像的影响。

“目前光声成像领域中的深度学习新方法,通常是将其他领域发展好的网络模型直接用于解决光声成像中的问题。如何改造其神经网络使它们更加贴近光声或其他成像技术的数理模型将成为深度学习新方法在生物医学成像领域应用的重要问题之一。”李娇说。

利用该深度学习新方法,团队成功重建出高空间分辨率的深层组织光学吸收系数定量分布图像。

这是首次应用定量光声层析成像深度学习新方法实现对活体深层组织光学吸收系数的“真实透视”成像。无标注真实数据情况下深度学习神经网络的成功应用也开拓了深度学习新方法在生物医学成像方向的发展空间。

瞭望站

瞄准智造新赛道 上海开年这么干

◎新华社记者 姚玉洁 龚雯 周蕊

能早则早、能快则快!新春伊始,抢占新赛道、培育新动能的上海步伐全力加速。从中芯国际千亿元投资规模的12英寸晶圆生产线,到上汽智己智能网联汽车,再到虚拟偶像“生活体验官”,上海正全力加速数字技术赋能实体经济的步伐。在GDP迈过4万亿元大关之后,“新赛道”和“智造力”正为上海积聚高质量发展的新势能。

赶前抓早,虎虎生风

春节前夕,上海设立了数十亿元规模的“国产汽车芯片专项基金”,突破卡脖子技术,缩短开发周期,为中国乃至全球汽车产业的创新发展贡献“上海智慧”;汇聚了国家级人工智能产业集群的上海市徐汇区,提出“探索成立元宇宙创新联盟”……

马不停蹄的还有仅用22个月实现“从零到一”、刷新“上海智造”速度的智己汽车。今年以来,首批智己L7 Beta体验版在全国多地的大街小巷巡游,预计2月底智己L7试驾驾车陆续到店,3月初将全面开启实车试驾。

这些项目是上海发展新技术、新业态、新模式的缩影。今年的上海两会,“强化新赛道布局 and 终端带动”被写入市政府工作报告;春节后首个市政府常务会议原则同意《上海城市数字化转型标准化建设实施方案》。“谁能占据经济数字化新赛道,谁就能更好掌握未来发展主动权”,成为这座经济中心城市的共识。

“智造力”就是“生产力”

刚刚过去的跨年营销中,一个穿着黑红裙白靴的身影让人眼前一亮,她是虚拟偶像Gina,也是百联股份“美好生活体验官”,通过探店Vlog在抖音、小红书、微信、微博等头部社交平台正式与消费者见面“带货”。春节期间,Gina还与冰雪运动“擦出火花”,点燃冬奥带来的冰雪消费。

专注“新赛道”和“终端带动”,终端产品正在带来全新的生产方式、消费方式和生活方式。

作为世界上最大的乳制品单体工厂,光明乳业华东中心工厂每天出品的牛奶可供数百万居民饮用,但工厂生产线上几乎看不到工人,只能看到许多机器控制的阀门和管道。通过这一智能工厂内的“智慧大脑”中央控制室,数字化操作人员仅通过中央控制室就能对牛奶“了如指掌”,工厂里5000个传感器、1万多个自动控制阀互联互通,保障工厂自动化运行。

据统计,2021年上海部分重点产业已建成智能工厂的生产效率平均提高20%,运营成本平均降低25%,综合能耗最多降低54%。

这种“智造力”正在向面广量大的中小企业输出、下沉。上海宾通智能科技有限公司致力于成为提供智能物流、柔性制造方案的“工厂大脑”。公司董事长龚超慧说:“我们正在和奉贤区东方美谷的美妆企业合作,帮助它们从接单管理、需求预测、订单排期、物料最优分配、生产实时追踪到数据回溯形成数字化闭环。中国有完善的供应链和工业体系,给中小企业的‘智造力’提升带来巨大的历史机遇。”

迈过4万亿,积聚新势能

强化新赛道布局,强化终端带动,上海为何反复对此划重点?“这是上海在GDP迈过4万亿元大关、站上发展新台阶之后,对发展动能、发展质量深度思考后的战略部署。”上海社科院党委书记权衡说,新供给创造新需求,上海发力新赛道,将有力带动经济引擎增长。

2021年上海建成2家国家级标杆性智能工厂、5家市级标杆性智能工厂、40家市级智能工厂。其中,被业界视为衡量智能制造水平的重要指标——机器人密度,上海重点产业达到383台/万名工人,部分企业如海立、新时达高达近1000台/万名工人,远超全国246台/万名工人的平均水平。2022年,上海有望成为全球首个发布机器人密度的城市。

“智造已刻进了上海发展的基因。”上海市经济和信息化委员会智能制造推进处处长韩大东说,随着上海以智能工厂建设推动经济数字化转型的不断深入,土地不再是强黏性,加上专业人才、先进技术、金融生态的优势,上海正在积聚新一轮数字驱动的“智造力”。

上海市经信委主任吴金城说,要紧抓新赛道、新动能,持续招引布局和培育做强一批数字经济、绿色低碳、元宇宙、智能终端等领域重大产业项目。既要着眼当前,聚焦“发热”和压舱石产业,又要立足长远,聚焦“发光”和引擎产业。



视觉中国供图