



上图:无人驾驶的物流配送车(资料图片)

清晨,北京的张先生被智能音箱轻柔的音乐唤醒。“小爱同学,今天天气怎么样?”他习惯性地问道。“今天北京天气晴转多云,气温18到25摄氏度,空气质量良,建议穿薄外套。”音箱立即给出了回答。这条看似简单的对话结果的背后,是气象卫星以及数万台地面气象设备和海洋浮标每秒钟产生的上亿条数据的支撑。而让机器能够听懂人类语言并作出恰当回应的能力,则来自人工智能技术的突破。

上海的某三甲医院,放射科李医生正在使用AI辅助诊断系统查看患者的CT影像。系统在0.3秒内就标记出了几个可疑的结节区域,并给出了91.7%的恶性概率评估。这个判断基于系统学习过的超过100万例肺部CT影像数据。李医生仔细核对着系统的建议,不禁感叹:“十年前,这样的技术还只存在于科幻电影里。”这就是我们正在经历的数字新时代——大数据与人工智能构成的“数字助手”已经进入我们生活的方方面面。从清晨的智能闹钟到通勤时的导航推荐,从工作邮件自动分类到晚间购物时的个性化推荐,这些技术正在悄然改变着我们的生活方式、工作模式乃至重塑整个社会结构。

大数据的特征

大数据是指规模巨大、类型多样且处理速度快的数据集,这些数据超出了传统数据库软件工具捕获、存储、管理和分析的能力。大数据的核心特征通常被概括为“4V”。

大量化(Volume):顾名思义,大数据的特征首先就体现为“大”,这个大是指数据的规模庞大。一般来说,存储一个汉字需要两个字,也就是2B;存储一张清晰度不是很高的照片需要几KB(千字节,1KB=1024B),存储一首歌曲需要几MB(兆字节,1MB=1024KB),存储一部高清电影需要几GB(千兆字节,1GB=1024MB),而大数据存储则是从TB(万亿字节,1TB=1024GB)级别起步,到PB(千万亿字节,1PB=1024TB)甚至EB(百亿亿字节,1EB=1024PB)级别。

快速化(Velocity):大数据的生成,需要及时快速地处理。在物联网时代,传感器、监控设备等实时产生数据流,需要即时处理才能发挥价值。比如一次简单的网购,在我们点击确认付款的瞬间,系统会完成库存核对、信用验证、物流匹配等多达十几个环节,而这只需要现实中不到1秒的时间。

多样化(Variety):数据的来源和格式是多种多样的。现代数据不再局限于结构化数据(如表格数据),还包括半结构化数据(如一份简历),以及非结构化数据(如文本、图像、视频等)。目前,绝大部分的数据将是非结构化的。我们可以用一个例子来说明,当你去医院体检时,你的身高体重和血压等数值就是结构化数据,CT影像就是半结构化数据,医生的医嘱就是非结构化数据。

价值密度低(Value):尽管大数据的体量巨大,但真正有用的信息可能只占很小一部分。比如一段几小时的监控录像,大部分画面可能都是正常的人流往来,真正有价值的可能只是其中几分钟的异常情况;一款热门App每天收集到的用户行为数据里,多数是常规操作,只有通过分析大量数据中隐藏的规律,才能找到用户的偏好,进而精准推荐。

大数据与人工智能

郭旭敏

人工智能的发展

人工智能(AI)是一门研究如何让计算机模拟人类智能的科学。它涉及多个领域,包括机器学习、深度学习、自然语言处理、计算机视觉等。通过算法和大量数据的训练,AI系统能够执行复杂的任务,如语音识别、图像分类、自动驾驶等。人工智能的发展跌宕起伏,经历了多次繁荣与寒冬,大致可分为以下几个阶段:

萌芽期(1950s~1960s)。1950年,英国计算机科学家艾伦·图灵发表《计算机器与智能》,提出“图灵测试”概念。1956年,达特茅斯会议正式确立“人工智能”这一术语,早期AI研究聚焦于符号逻辑和通用问题求解。这一时期出现了首个人工智能程序Logic Theorist(1956)和首款聊天机器人ELIZA(1966)。

知识工程与专家系统(1970s~1980s)。研究者意识到单纯依靠通用推理难以解决复杂问题,转而开发基于专业领域知识的专家系统。医学诊断(MYCIN)、化学分析(DENDRAL)等专业系统的出现,展示了AI在这方面的实用价值。

机器学习兴起(1990s~2000s)。随着计算能力提升和数据量增加,统计学习方法开始取代基于规则的系统。支持向量机(SVM)、随机森林等算法在诸多任务上表现优异。1997年,IBM深蓝击败国际象棋冠军卡斯帕罗夫,展示了AI的潜力。这一时期也是搜索引擎和推荐系统快速发展的阶段。

深度学习革命(2010年以来)。神经网络研究在经历低谷后迎来复兴。2012年,AlexNet在ImageNet竞赛中大幅领先传统方法,引爆深度学习热潮。随后,深度学习在计算机视觉、语音识别、自然语言处理等领域取得突破性进展。2016年,AlphaGo击败围棋冠军李世石,引发全球关注。Transformer架构(2017)的出现,推动了大型语言模型的发展。

基础模型时代(2020年以来)。以ChatGPT、DeepSeek等为代表的大规模预训练模型展现出强大的泛化能力和新兴特性。多模态模型

可以同时处理文本、图像、音频等多种数据。AI开始应用于科学发现、内容创作等复杂领域。同时,AI伦理、安全、治理等问题也受到前所未有的关注。

大数据与人工智能

大数据与人工智能之间存在着深刻的协同效应和相互依赖关系。

数据是AI的“燃料”:现代AI尤其是机器学习算法依赖大量数据进行训练。数据规模和质量直接影响AI系统的性能和准确性。例如,深度学习模型通常需要数十亿个训练样本才能达到理想效果。

AI释放数据价值:面对海量数据,传统分析方法已力不从心。AI技术特别是机器学习算法能够高效处理和分析大数据,从中提取有价值的信息。AI可以视为从数据中提取知识的“工具箱”。

相互促进的良性循环:更好的AI技术可以处理更多样、更复杂的数据,而更丰富的数据又能训练出更强大的AI模型。这种正反馈循环推动着两个领域的快速发展。

基础设施的共享:大数据处理框架为AI提供了可扩展的计算平台,而AI的优化算法又提高了大数据处理的效率和智能化程度。

人工智能旨在创造能够模拟、延伸和扩展人类智能的理论、方法、技术及应用系统。人工智能的核心目标是使机器能够胜任通常需要人类智能才能完成的复杂任务。根据能力水平,AI可分为三类。

弱人工智能(Narrow AI):专注于执行特定任务的人工智能系统。这类AI在特定领域表现出色,但不具备真正的理解或意识。现在生活中我们常接触的AI应用都属于此类,如语音助手、推荐系统、图像识别软件等。

强人工智能(General AI):理论上具有与人类相当的综合认知能力,可以自主思考、学习并解决各种问题的AI。这类AI目前尚未完全实现,是许多研究者的长期目标。

超级人工智能(Super AI):在几乎所有领域都远超人类认知能力的AI。这种AI目前只存在于科幻作品和理论讨论中,其潜在影响引发了广泛伦理和安全讨论。

大数据与人工智能技术已进入到各行各业,深刻地改变着传统业务模式和社会运行方式。

智能交通管理 现代城市智能交通管理通过摄像头、传感器、GPS设备等基础设施实时采集海量交通数据,包括车辆位置、道路状况和交通流量等信息。这些数据经由大数据平台整合后,由AI算法进行深度分析与预测,发展出了多项创新应用:杭州“城市大脑”利用AI动态调节红绿灯时长,使局部区域通行效率显著提升15%以上;基于计算机视觉技术的事后自动识别系统大幅缩短了应急响应时间;通过挖掘历史停车数据,系统能精准预测需求热点并引导车辆合理分流;同时依托客流大数据分析,公共交通部门可科学优化线路规划和班次安排,从而全面提升城市交通运行效率。

公共安全与城市治理 现代城市公共安全与治理体系正通过AI技术实现全面升级,其中犯罪预测系统通过整合历史案件数据、气象条件和节假日特征等信息,预判犯罪高发区域,为警方提供巡逻路线优化和警力调配决策支持。智慧“天眼”已发展成为城市治理的核心基础设施,像上海“一网统管”平台为例,其接入的数百万摄像头,既能快速定位失踪人员轨迹,也可实时识别公共场所的异常行为模式。在市政设施管理领域,物联网传感器持续采集桥梁振动、管道压力等结构参数,通过大数据分析提前发现潜在故障隐患,实现从被动抢修到预测性维护。

医学影像分析 AI技术在医学影像分析领域取得了显著进展,其能力在放射影像解读和病理切片分析等多个方面已经达到甚至超越了人类专家的水平。应用成果包括Google DeepMind开发的乳腺癌检测系统,该系统在多项关键性能指标上都展现出了超越专业放射科医生的精准度。同样,腾讯推出的觅影平台也展示了强大的应用价值,能够有效地对食管癌、糖尿病视网膜病变等多种疾病进行早期筛查,为疾病的及时干预提供了有力支持。

商业与工业场景 在智能供应链管理方面,通过整合销售数据、天气信息和社交媒体趋势等多维数据,AI能够更精准地预测产品需求;同时借助动态库存优化系统,企业可以实时调整库存水平,有效减少缺货和库存积压问题;而基于实时交通数据的智能物流规划系统,如京东的智能配送系统,则大幅提升了配送效率。在个性化购物体验领域,推荐系统发挥着关键作用,亚马逊平台上35%的销售额都来自个性化推荐;视觉搜索功能让消费者可以通过拍照或上传图片来寻找相似商品;AR技术实现的虚拟试衣间则让顾客能够在“线”试穿“服装”,大大提升了购物体验的真实感和便利性。

这些应用场景只是大数据与AI改变世界的冰山一角。随着技术不断进步,其应用边界将持续扩展,创造更多我们难以预见的可能性。理解这些实际应用有助于我们把握数字化转型带来的机遇。

社会影响与伦理挑战

随着大数据技术和人工智能技术的深入,产生了一系列亟待解决的伦理和社会问题。

隐私安全:近些年的数据显示,医疗机构数据泄露事件同比增加40%,其中83%源于第三方合作方数据滥用;生物识别技术引发“透明人”困境,人脸识别诉讼年增120%,涉及社区门禁、



智能家居(AI生成)

考勤等多种场景;通过“性别+出生日期+邮编”组合可识别87%匿名用户。

算法偏见:招聘算法存在性别偏见,同等条件下女性收到高薪职位推荐频率低40%;外卖算法将“准时率”权重设为85%,间接导致骑手违规率上升300%。

职业替代:智能工厂使单个车间用工量从120人减至20人;会计、法律文书等岗位自动化率已达45%(2025年人社部数据);网约车司机每日接受算法调度指令超200次,决策自主权下降70%。

偏见放大:产生信息茧房,推荐阅读系统使用户接受多样性观点的概率下降67%,极易引发社会对立情绪。

文化同质化风险:据调查研究显示,目前全球96%的新闻推荐算法优先推送英语内容,直接导致其他语种内容可见度下降40%。

大数据与人工智能的未来发展路径将深刻影响人类社会走向。面对这一重大历史机遇,我们需要智慧和勇气,既要充分发挥技术潜力以解决人类面临的重大挑战,又要未雨绸缪防范潜在风险,引导技术向增进人类福祉的方向发展。这不仅是科学家和工程师的责任,也是每个人应当关注和参与的公共议题。

AI是否会取代人类?答案并非简单的“是”或“否”,而是一个复杂的问题。目前主流观点是:AI不会完全取代人类,但它将深刻改变劳动力市场的结构,淘汰部分职业,同时创造新的机会。

AI和自动化技术最适合替代重复性高、规则明确、低创造性的工作。可能逐渐被取代的领域有以下几种:

制造业重复度高的流水线工作;工业机器人已广泛应用于汽车、电子等制造业,完成焊接、组装、包装等任务,中国制造业的自动化率预计2030年达50%。

基础数据管理与行政工作:AI可自动处理发票录入、数据整理、基础客服等任务,初级文员、数据录入员等岗位需求将大幅减少。

运输与物流:无人卡车已在干线物流试运行,预计2030年普及率超30%,配送机器人可完成“最后一公里”送货,卡车司机、快递员等职业面临转型压力。

尽管AI能力强大,但在以下领域仍难以替代人类。

创造性工作:AI可生成图像或音乐,但创意构思、情感表达仍需人类主导;AI可辅助写作,但深度思考、文化洞察是人类专长。

复杂决策与情感交互:AI能分析数据,但团队激励、战略判断仍需依赖人类经验;共情能力(如教师对学生个性化指导)无法被机器完全复制。

灵活应变与非标准化任务:消防员、医生等需在多变环境中快速反应;厨师、高级钟表匠等依赖独特技艺。

AI的普及将催生大量新职业,主要集中的领域有以下几种:

AI训练师与数据标注专家:训练AI模型(如标注图像、优化对话系统)确保数据质量,减少算法偏见,预计到2030年时,全球岗位将超500万个。

AI伦理与合规专家:制定AI使用规范(如隐私保护、公平性审查)评估算法风险(如自动驾驶的道德决策)。

人机协作设计师:优化AI与人类的交互体验(如语音助手、智能家居)设计“增强智能”工具,辅助人类决策。

智能系统维护员:维护机器人、自动驾驶车辆等硬件设备,调试AI系统故障,确保稳定运行。

应对AI挑战需个人与社会协同努力。个人应重点培养创造力、批判性思维和高情商等AI难以替代的能力。社会层面需完善就业保障体系,提供职业培训等支持,共同构建人机协作的可持续发展生态。

AI不会完全取代人类,而是成为生产力的倍增器。未来的工作模式将是人类与AI协作,人类负责创意、决策与情感交互,AI处理重复性任务。面对变革,我们应当主动适应,社会需完善教育与保障体系,共同构建人类与AI协作的新时代。

人工智能的发展需坚守“以人为本、智能向善”原则,通过技术创新与伦理规则的深度融合,确保技术红利惠及全人类,迈向安全、公平、可持续的智能文明。

拥抱智能时代

从结绳记事到数字文明,人类始终在寻找更有效的信息处理方式。大数据与人工智能代表了这一探索的最新阶段,它们正在重塑我们理解世界和解决问题的方式。

技术变革的加速度方面:大数据与AI的发展速度远超以往任何技术革命。从1956年AI概念首次提出到今天DeepSeek、ChatGPT大模型的广泛应用,从GB级数据到现在的ZB级数据时代,技术进步呈指数级增长。这种加速变化要求我们的学习方式和思维模式也必须与时俱进。

普惠价值的实现方面:当技术门槛降低,大数据与AI将成为像电力一样的基础设施,赋能每个个体和小微企业。农民通过手机获取精准农业建议,小店主利用数据分析优化库存,学生借助AI辅导获得个性化学习——技术平权将释放巨大创造力。

人机协作的新范式方面:与其担心AI取代人类工作,不如关注如何与AI协同合作。医生使用诊断辅助系统提高准确率,设计师利用生成工具拓展创意边界,研究者借助文献分析系统追踪科学前沿——人机协作将创造“1+1>2”的价值。

持续演进的数字生态方面:我们正在构建一个数据驱动、智能增强的数字生态系统。这个系统中,数据如同土壤,算法如同作物,在不同应用场景开花结果。保持这个生态的健康平衡需要技术创新、制度设计和人文关怀的协同。

共同参与方面:技术发展方向最终取决于社会选择。我们应当了解基本概念,消除非理性恐惧或期待;关注技术应用的伦理影响和社会后果;参与公共讨论,表达对技术发展的期望和担忧;终身学习,适应技术带来的变化。

大数据与人工智能不是遥远的未来,而是正在发生的现实。从医疗诊断到天气预测,从教育个性化到艺术创作,这些技术正在扩展人类能力的边界。面对这一深刻变革,被动适应不如主动参与。

正如工业革命不是终结于蒸汽机的发明,而是开启了持续创新的长河,我们今天也正处于一场伟大变革的开端。大数据与人工智能的潜力远未完全释放,其最终形态将由我们这代人的选择和行动共同塑造。理解它、思考它、引导它,是我们共同的责任和机遇。

讲座地点:山西青年职业学院

主讲人:郭旭敏

时间:2025年4月17日



右图:物流智能仓储(资料图片)

