

科技前沿

人工智能，从深度学习到全面赋能

龚 克

近10年来，以深度学习为主要技术路线的人工智能(AI)，从判别(如图像、文字、语音识别等)到生成(如文本、图像、视频、程序等)接连取得突破，DeepSeek、ChatGPT、Sora等应用展示多种技术可能。作为革命性通用技术，人工智能正加速融入人们的日常生活、生产工作和学习研究之中，推动人类社会信息化进程迈向智能化的新阶段。

发展新阶段：从判别到生成，从单模态到多模态，从专用到通用

人工智能作为近年来科技领域最热门的话题之一，其实并不是一项新技术。早在1956年，人工智能的概念就在美国达特茅斯学院召开的夏季研讨会上被提出。此后经历半个多世纪发展，都未能实现真正的实用化。2012年，加拿大多伦多大学教授杰弗里·欣顿和他的学生基于机器学习提出的深度学习神经网络模型，在图像识别挑战赛上取得重大突破，激发AI领域的又一次发展热潮，推动语音识别、图像识别以及自然语言处理等多个领域快速进步。

AI技术不断迭代升级。特别是2022年前后，生成式人工智能迎来全面爆发。2022年11月，美国OpenAI公司推出ChatGPT，以更大规模的模型实现了更强的内容生成(AIGC)功能。2024年2月，OpenAI公司推出的文生视频大模型Sora，推动生成式人工智能从只能理解单一形式信息(比如文字、图片、音频等)的单模态，迈向可以同时理解多种形式信息(比如图片和文字)的多模态。与此同时，全球的AI大模型竞相超越又各有特长，国外谷歌的双子座(Gemini)、安索皮克(Anthropic)的克劳德(Claude)、元公司的Llama、中国的文心一言、智谱清言、豆包、书生·浦语、星火认知、360智脑、通义千问、混元、Kimi、零一万物等，形成百花竞放的局面。

2025年1月，我国人工智能企业深度求索(DeepSeek)推出新模型DeepSeek—R1，凭借技术创新和商业化潜力引发全球瞩目。这不仅是因为该模型在自然语言处理方面表现卓越，更为重要的是其在训练方法上实现了创新，成功降低了同类产品对算力资源的大量需求，显著降低了使用成本。得益于此，用户可将R1部署在个人本地设备，实现更为安全且个性化的服务。尤为重要的是，DeepSeek的开源模式为全球创新发展带来新机遇，并有可能带动全行业在未来形成以开源为主流的发展路线。

从判别到生成，从单模态到多模态，从专用到通用，是智能化发展新阶段的重要标志。人工智能以自然语言为基础的表达方式，正从方方面面渗透进大众生活。从手机应用、智能音箱等智能助手，到智能网联车和智能机器人等具身智能，再到能提供个性化服务的智能体以及更多的AI原生产品和服务，正加速向我们走来。

链接

AI同学，请回答

网友问：什么是深度学习？可以通俗易懂地介绍一下吗？

DeepSeek：深度学习(DNN)是一种模仿人脑神经结构的机器学习技术，专门处理复杂数据。你可以把它想象成一个“多层级加工厂”：比如识别猫的照片，第一层先看像素点，第二层组合成线条，第三层拼出猫耳、胡须等特征，最后一层综合判断这是猫。每层由大量“神经元”(数学计算单元)组成，层与层之间通过可调节的“连接强度”传递信息。

它的核心能力在于自动学习特征。传统算法需要人工设计规则(比如，先找边缘、再识别形状)，而DNN能通过海量数据自己发现规律。训练时就像教小孩认字：先猜结果，对比正确答案，再反

核心驱动力：计算机能够从数据中学习规律并做出预测或决策

机器学习是指通过数据训练模型，使计算机能够从数据中学习规律并做出预测或决策。作为当下AI发展的主流技术路线，这条路线让机器通过学习而不是靠预设规则来获得智能，体现了学习是智能形成和发展的本质。可以说，生物智能也好、机器智能也好，都离不开学习，所有的成功和突破都源自学习。“生而知之”毕竟是极少数天才的特质，“学而知之”才适用于大多数人。人类在发展人工智能时也采用了学习的路线。

机器学习主要通过神经网络实现。神经网络作为机器学习的一种具体模型，就是通过模拟人脑的结构和功能来处理数据。人类大脑是由神经元和突触构成，神经元通过突触传递信号，实现信息处理、学习和记忆等功能。人工神经网络模拟这一结构，利用集成电路芯片组成的电子计算系统，采用人工神经元和连接(即模型参数)代替生物神经元和突触，通过输入数据、调整权重来学习和预测，构成了AI的物质底座。

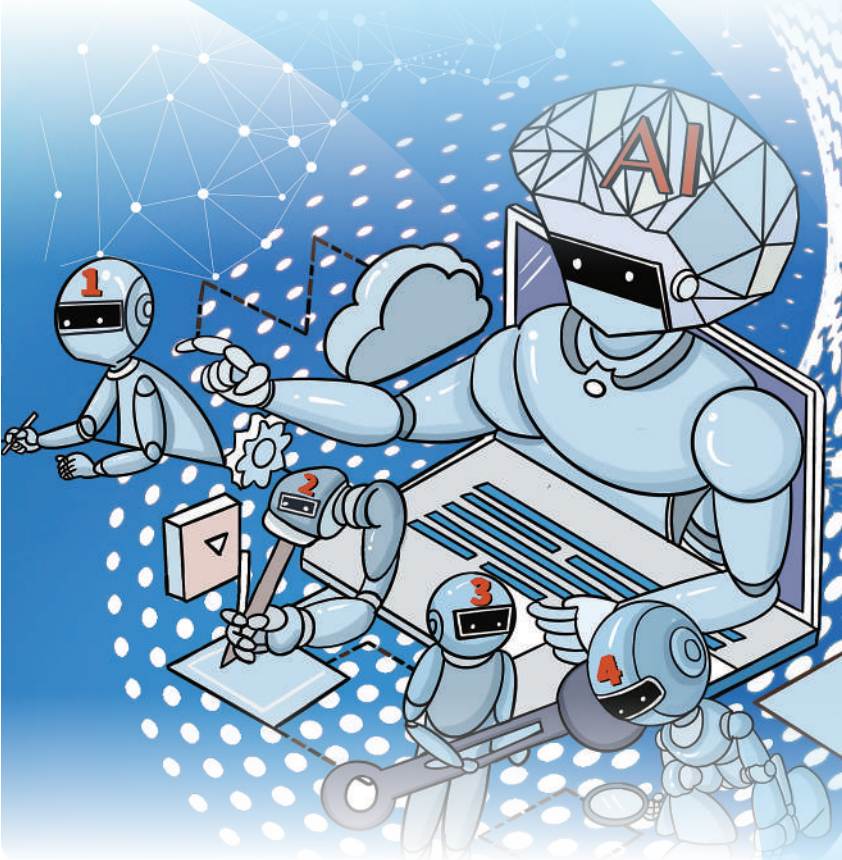
正如人脑层数越多，处理复杂信息的能力就越强，神经网络也同样需要更多的层数，这就需要借助深度神经网络。以深度神经网络构建的深度学习系统即为AI“模型”，这类模型的参数越多，就意味着模型越大，性能越强。例如，GPT—3模型实现了千亿量级参数，GPT—4模型参数甚至达万亿量级。

简而言之，机器学习是人工智能的核心驱动力，它通过从海量数据中提取有价值的信息，帮助人类做出预测、判断和生成所需的内容，从而提升任务处理效率，将人力从繁重、重复、低级和危险的劳动中解放出来。2024年的诺贝尔物理学奖颁给了约翰·霍普菲尔德和杰弗里·欣顿，以表彰他们通过人工神经网络实现机器学习的基础性发现和发明。展望未来，机器学习通过不断发展，必将为人类社会带来更大惊喜。

更广泛运用：AI智能体可望成为今年发展应用亮点

2025年将是AI走向更广和更深应用的关键之年，AI智能体(或称AI代理)可望成为今年AI发展应用的亮点。AI智能体可以说是人工智能技术的集大成者，具备自主性、感知能力、决策能力和行动能力等显著优势。它能够通过理解交办的任务、感知相应的执行环境，进而自主调用相应的智能模型功能、采取相应的行动，可以广泛应用于工业、医疗、交通、金融和家庭等领域。每个用户可能有多个AI智能体帮助解决各种需求，它们有的像是秘书，有的像是工匠，有的像是医护，有的像是采购员，有的像是清洁工。

例如，某公司计划设计一个活动邀请系统，在将活动时间、地点和主题等信息准确传达后，AI可自动生成一个智能体“小组长”，负责整体规划和调度。小组长可以召唤多个执行不同具体任务的智能体，1号智能体负责活动邀请函的策划，2号智能体负责设计



邀请函中的图像及文字，3号智能体负责代码制作和部署网页，4号智能体负责检验差错，他们就像是一个“工作专班”，成为完成个性化工作的“群体智能”。

通过AI智能体，各种AI技术将会被更方便、更广泛、更有效地应用起来，将在多个领域引发根本性变革。同时，快速发展的“科研智能”正在形成科学发展的新范式，为科技这一第一生产力的发展赋能，意义深刻而长远。在这个日新月异的智能时代，学会应用AI技术也将成为多数人的必备技能，人们需要学会以有效而适当的方式运用AI助益工作和生活。

“两岸猿声啼不住，轻舟已过万重山。”随着AI在生命科学、物质科学、数学、工程技术等领域的应用日益广泛和深入，可以预见，正如历史上机械动力和电气技术极大拓展了人类的体力劳动能力一样，AI正在加速拓展人类的脑力劳动能力，推动生产力实现又一次质的飞跃，并对人类社会的发展进步产生全面而深远的影响。

在AI先进生产力驱动的社会智能化历史进程面前，人类要把握和主导自身的命运，就必须提升自身的智能化能力，学会驾驭AI。各方应努力让AI行驶在促进经济繁荣 and 环境保护、保护公共利益和个人隐私、遵守法律法规和伦理良俗、促进公平正义与包容和谐的轨道上，确保AI造福全人类的可持续发展。

(作者为中国新一代人工智能发展战略研究院执行院长)



《讽刺与幽默》供图



不过，大语言模型并不是真正“理解”文字，它更像是在模仿人类的语言模式。它没有感情，也不会主动思考，只是根据输入的内容给出最合适的回应。所以，有时候它的答案可能看起来很合理，但其实是错的，这就需要我们去判断和核实。

专家点评：大语言模型不仅网络规模庞大，还采用一个重要的技术即“词嵌入”技术。该技术将自然语言的“词”嵌入到多维的语义空间，表达为“语义矢量”，从而可以对“词”进行计算，由此可以跨越语言的形式(如英语、汉语、文本、图像等)，生成各方面内容。大语言模型的突破推动了人工智能几乎所有(应用语言)领域的应用，但也面临生成偏差、算力消耗、伦理风险等挑战，需结合人类监督与技术优化以实现安全可控的发展。

(专家为中国新一代人工智能发展战略研究院执行院长龚克)

网友问：我们应该如何理解大语言模型这个概念？别太复杂。

通义千问：大语言模型可以理解为一个“超级聪明的文字助手”，它通过学习海量的文本数据，掌握人类语言的规律和知识。你可以把它想象成一个读过无数书籍、文章、网页的“大脑”，它知道怎么回答问题、写文章、讲故事，甚至还能编程、翻译、聊天。

它的核心能力是预测下一个词。比如你输入一句话的开头，它会根据上下文猜测接下来最可能说什么。通过这种方式，它能生成连贯、自然的文字。而且，因为它学了很多知识，所以不仅能回答简单问题，还能处理复杂的任务，比如分析逻辑、总结长文或者模仿某种写作风格。

科技大观

近年来，生成式人工智能热潮席卷全球，给世界带来巨大机遇的同时，也带来难以预知的各种风险和复杂挑战。人工智能全球治理的讨论与合作在多个层面展开，逐渐扩展至以法律规制和技术标准为核心的制度建构。

人工智能治理的议题涵盖广泛，涉及人工智能技术开发、部署、应用中的伦理道德与行为规范，以及人工智能带来的产业结构变化、社会权力结构变化、地缘政治影响、人类生存风险等诸多方面。人工智能的技术特性，使其既具有互联网服务应用的广泛性、赋能性，又具有作为基础设施对国家安全的重要影响。这决定了其治理体系在设计之初就需要兼顾促进创新的发展目标和防范风险的安全需求。

在人工智能的国际治理中，近年来初步形成了多个互相竞争合作的板块，其中以一些人工智能安全峰会与联合国框架下的国际合作最为引人关注。2023年11月，首届全球人工智能安全峰会在英国布莱奇利郡举行，确立了人工智能安全治理的全球共识，包括中国在内的28国及欧盟签署宣言。2025年2月，在法国巴黎举办了人工智能行动峰会，包括法国、中国、印度、欧盟在内的多个国家和国际组织共同签署了《关于发展包容、可持续的人工智能造福人类与地球的声明》。联合国在其框架下积极推进人工智能治理全球合作，尤其是具有包容性的、普惠的人工智能全球治理。2023年10月，联合国秘书长古特雷斯宣布成立人工智能高级别咨询机构，并形成《以人为本的人工智能治理》报告。2024年9月，联合国举办未来峰会，通过《全球数字契约》，倡导开启人工智能治理的全球对话，推动人工智能技术负责任发展。

中国积极参与和引领人工智能全球治理进程，向世界贡献了重要的人工智能治理观。一是以人为本促进发展。在我国既有人工智能治理实践中，坚持发展与安全并重，强调以人为本、智能向善，让人工智能技术真正助力全球发展事业，赋能人类美好生活。中国发布的《全球人工智能治理倡议》《人工智能全球治理上海宣言》也充分体现了这一立场。二是加强能力建设，弥合数字鸿沟。我国积极加强与全球南方国家的人工智能能力建设合作。去年7月，联合国大会协商一致通过中国主提的加强人工智能能力建设国际合作决议，140多国参加决议联署。中国—金砖国家人工智能发展与合作中心也于去年成立。这些举措旨在缩小国家间及国家内部在人工智能及其他数字领域存在的差距，针对发展中国家在人工智能设计、开发、部署和使用上面临的独特挑战，提供了加强南北合作、南南合作和三方合作的破局之道。三是推动人工智能发展共治共享。人工智能全球治理应坚持广泛参与、协商一致原则，在充分尊重各国政策和实践差异性的基础上，推动形成广泛国际共识。

整体来看，目前全球人工智能治理中，各国已就伦理道德和原则开展了大量讨论，但相关法律法规、国际条约和治理机构仍处于起步阶段，需将不断探索的各类举措汇聚成以国际法和可持续发展目标为基础、连贯一致、可互操作的整体设计，并随着技术发展进步不断作出调整。去年在上海举行的2024世界人工智能大会暨人工智能全球治理高级别会议将会议主题确定为“以共商促共享 以善治促善智”，体现了《全球人工智能治理倡议》的核心要义，反映了国际社会的共同呼声。中方将一如既往支持各国在人工智能治理中秉持共同的安全观，协力促进人工智能治理，增强发展中国家在人工智能全球治理中的代表性和发言权，推动人工智能更好提升人类福祉。

(作者为中国政法大学数据法治研究院教授、联合国人工智能高级别咨询机构中方专家)

人工智能助力新模型实时监测全球珊瑚礁状况

据新华社堪培拉电 (记者章建华、李惠子)南澳大利亚大学等机构的研究人员借助人工智能，开发出一个可对全球珊瑚礁进行实时监测的模型，以实现早期干预，减缓气候变化对珊瑚礁的危害。该研究结果日前发表在新一期学术期刊《电子学》上。

南澳大利亚大学等机构的研究人员借助人工智能，并结合遥感技术、机器学习和地理信息系统等技术开发出一个新模型，可将影响珊瑚礁的所有因素整合起来，以进行实时全球监测。相关水下视频和照片、卫星图像、文本文件和时间传感器数据等都将被提取并显示在统一的中央仪表盘，通过实时集中这些信息生成有助于保护珊瑚礁工作的预测模型，从而实现更早干预。

研究人员用AI设计出具有天然酶关键特征的全新酶

据新华社洛杉矶电 美国研究人员近日在《科学》杂志上发表论文说，他们利用人工智能(AI)设计出了具有天然酶关键特征、可介导多步反应的全新酶。业内专家评价其为“酶工程领域的一个里程碑”。

研究人员说，他们此次借助多种AI工具，设计出了一种可介导多步反应的丝氨酸水解酶，这种酶的催化效率比之前设计的类似酶高出6万倍。丝氨酸水解酶是一类在催化水解反应中起关键作用的酶，天然的丝氨酸水解酶参与许多生物过程，包括消化、脂肪代谢、凝血等。

研究人员表示，虽然此次只是实现了原理验证，新设计的丝氨酸水解酶在效率上仍不及天然的丝氨酸水解酶，但他们希望进一步调整酶的结构以提高催化效率，使这项技术更接近实用，比如设计出用于分解塑料的酶等。