

2024 年诺贝尔科学奖中的未来产业场景

编者按

2024 年诺贝尔物理学奖和化学奖都颁给了 AI 相关领域的科学家,标志着 AI 与学科交叉的新时代的到来。AI 已不仅仅是加速科学研究的工具,而是推动科学范式转变的关键力量。AI 能够处理海量数据、模拟复杂系统,并生成新的科学假设,这种数据驱动的科研模式将进一步扩展至其他科学领域,彻底颠覆人们理解自然世界的方式。

诺贝尔化学奖

我的导师戴维·贝克执著于“蛋白质设计革命”

▶ 本报特约作者 王初

当地时间 10 月 9 日,瑞典皇家科学院宣布将 2024 年诺贝尔化学奖颁给 3 位科学家,其中两位是 Foldit 的创始人——英国科学家德米斯·哈萨比斯(Demis Hassabis)和美国科学家约翰·江珀(John M. Jumper),表彰他们在蛋白质结构预测方面取得的成就;另外一位是我的导师——美国华盛顿大学生物化学家戴维·贝克(David Baker),表彰他在计算蛋白质设计方面的贡献。

我于 2001 年来到华盛顿大学生物化学系攻读博士学位,机缘巧合下通过实验室的轮转来到戴维·贝克实验室,一直到 2008 年博士后项目结束。在与戴维的接触过程中,我感觉到他是一个天生的非常纯粹的科学家,并且他把自己的全部精力都倾注于科学研究,总会有很多独到的想法,并将这些想法付诸实施。

因何设计蛋白质

戴维执著于“蛋白质设计革命”(protein-design revolution)。他说,这场革命与人类历史上发生过的其他科技革命类似:学会控制金属,让人类走出了石器时代;在工业革命中,人们学会了用蒸汽机来操控机器;而在之后,数字革命又让人们能在计算机中存储信息。现在,通过“蛋白质设计革命”,人们将学会用前所未有的方式操控生物分子。“我们可以用蛋白质制造出新的药物、疫苗、疾病疗法甚至新的材料。”

改变新药发现和开发模式

蛋白质是生命活动的直接执行者,其结构与功能由氨基酸一级序列所决定。上个世纪 90 年代,贝克研究组开始进行 Rosetta 软件包的开发,尝试对蛋白质结构进行计算模拟与预测。该方法的基本策略是,通过 Rosetta 搜索所有已知蛋白质结构的数据库,并寻找与目标蛋白序列相似的蛋白质短片段。利用蛋白质能量分布的基础知识,Rosetta 再对这些片段进行优化获得完整蛋白质。

1998 年,Rosetta 在蛋白质结构预测双盲竞赛 CASP 中首次亮相,与其他参赛者相比,它表现出非常好的预测效果。这一成功导致了贝克团队产生了一个新的想法——与在 Rosetta 中输入氨基酸序列并输出蛋白质结构相比,他们可以调用 Rosetta 的反向功能,即输入所需的蛋白质结构并获得其氨基酸序列的建议,这将使他们能够创建全新的蛋白质结构。如果能实现从开始设计蛋白质结构的功能,这意味着人们可以创造出具有各种功能的全新的生命活动执行者,并将其应用于生物医学研究的方方面面。

构建全新蛋白质“从头设计”

构建全新蛋白质的领域被称为“从头设计”(de novo design)。针对该目标,贝克团队绘制了一种当时在数据库中不存在的计算全新蛋白质结构,然后让 Rosetta 通过能量最小化计算氨基酸序列可以折叠成这种结构,并在此基础上继续优化改进,直到完成从氨基酸序列到蛋白质结构再到氨基酸序列的闭环验证。贝克研究小组将建议的氨基酸序列的基因引入细菌中,通过细菌表达产生所需的

蛋白质并后续纯化。然后他们使用 X 射线晶体学确定蛋白质结构。结果表明,实验解析的结构与 Rosetta 软件设计的结构完全一致。由于该蛋白是由经 7 轮设计产生,研究人员将这个蛋白质命名为 Top7。

Top7 的成功,对从事蛋白质设计的研究人员来说是一个惊喜。那些以前设计的蛋白质的人只能模拟现有结构。Top7 的独特结构在自然界中不存在。此外,由 93 个氨基酸组成的蛋白质,比以前使用从头开始设计方法产生的任何蛋白质都要大,而且由于经过多轮的迭代和优化,这个蛋白也非同等大小的天然蛋白质更加稳定。

贝克课题组于 2003 年发表了这项发现,这是令人惊叹的发展的第一步。鉴于蛋白质的结构可以决定其性质和功能,这为创建具有新功能的定制蛋白质奠定了基础。在他们之前的蛋白质设计和工程化研究中,研究人员大多是通过调整现有的蛋白质,以便它们能够产生新的功能(例如,分解有害物质或作为化学制造行业中的工具)。然而,天然蛋白质的范围有限,为了增加获得具有全新功能的蛋白质的潜力,贝克研究小组在 Top7 成功后,希望从头开始创建它们。正如戴维所说,“如果你想建造一架飞机,你不应该从修改一只鸟开始;相反,你应该了解空气动力学的第一原理,并根据这些原理建造飞行器。”

改变新药发现和开发模式

在过去的 20 多年里,贝克实验室涌现出许多令人瞩目的蛋白质创造,他还发布了 Rosetta 的代码,因此全球的研究团体都在继续开发该软件,寻找新的应用领域。

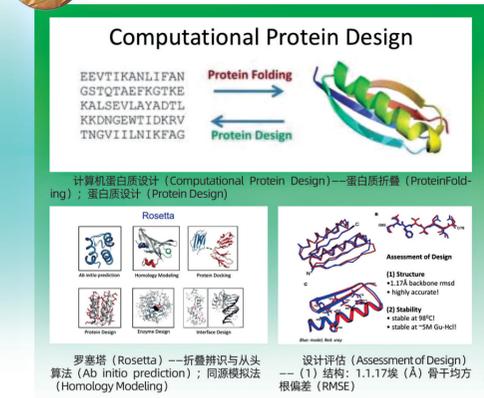
基于在蛋白质设计领域的深厚经验积累和深度学习算法带来的计算模式突破,贝克实验室近年在蛋白质设计领域的突破可以用日新月异来形容。从开发出一秒设计全新蛋白的工具,到设计可靶向任意蛋白的抗体,他的团队一次又一次地颠覆蛋白质设计的规则。这些在基础科学领域的突破也在通过初创公司的创建改变新药发现和开发的模式。戴维参与创建了超过 10 家生物技术公司,其中今年浮出水面的 Xaira Therapeutics 公司由 ARCH Venture Partners 和 Foresite Labs 共同孵化,获得启动轮融资超过 10 亿美元。

Rosetta 软件是基于蛋白质的生物物理特征,根据氨基酸序列计算预测蛋白质的三维结构。然而,预测蛋白质三维结构需要大量的计算力和时间,戴维在哲学和社会学方面的经验让他想到了利用神经网络来解决这个问题。他说:“我的理想是全球各地的人们都可以同心协力,为科学和全球健康做出贡献。”2004 年,他的实验室开发出的名为 Rosetta@home 的软件,使得任何人都可以在自己家里的计算机上下载这款软件,为解析某些蛋白结构的运算出一份力。在不到 4 年的时间里,接近 20 万志愿者在自己的计算机上安装了这款软件,为蛋白质结构预测做出贡献。

(作者:王初,博士,北京大学化学与分子工程学院教授)



诺贝尔化学奖



诺贝尔物理学奖

▶ 本报特约作者 唐乾元

物理界炸裂:这更像是计算机奖、数学奖

▶ 本报特约作者 唐乾元

神经网络拥有独特模拟能力

当地时间 10 月 8 日,瑞典皇家科学院宣布将 2024 年诺贝尔物理学奖授予美国 and 加拿大科学家约翰·霍普菲尔德(John Hopfield)和本弗里·辛顿(Geoffrey Hinton),以表彰他们在人工神经网络和深度学习领域的开创性贡献。他们的工作展示了如何从统计物理学中的基础概念出发,利用能量最小化、概率建模和优化等方法帮助计算机通过可计算的方式实现学习和记忆。他们开创性工作的科学价值在于:不仅为人工神经网络提供了理论支撑,还让深度学习成为解决现实复杂问题的强大工具,比如图像识别和自然语言处理。

为计算神经科学提供了重要的理论工具

约翰·霍普菲尔德的贡献可以形象地理解为给机器学习设计了一种“联想记忆”的能力。就像人们看到模糊破旧的老照片,也能通过记忆把它拼凑完整一样,霍普菲尔德的网络能够通过不断调整自身状态重现原始图像。这种网络模拟模拟了物理系统趋向最低能量状态的过程,类似于在

神经网络拥有独特模拟能力

2024 年诺贝尔物理学奖授予美国 and 加拿大科学家约翰·霍普菲尔德(John Hopfield)和本弗里·辛顿(Geoffrey Hinton),以表彰他们使用人工神经网络实现机器学习的历史性发现和发明。这是诺贝尔物理学奖历史上首次跨越传统边界,认可了计算机科学的重要进展。

神经网络拥有独特模拟能力

人工神经网络是一种模拟人脑神经元之间连接的计算模型,通过一个基于统计学的学习方法,在外界信息的基础上改变计算模型内部结构,实现对模型的优化。作为生物学、物理学与计算机科学领域的伟大结晶,人工神经网络在挖掘客观世界隐藏特征及其组合,进而实现对客观世界的模拟方面拥有独特能力。已在模式识别、智能机器人、自动控制、预测估计、生物、医学、经济、交通等领域辅助人类决策,成功地解决了许多现代计算机难以解决的现实问题。

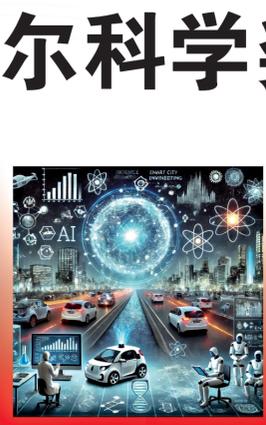
约翰·霍普菲尔德创造性地提出了可以由“开”和“关”的神经元组成的神经网络,即霍普菲尔德网络,它是一个模拟人类记忆的模型,通过用与物理学中自旋系统能量等同的一个概念描述网络的总体状态,并通过结合所有节点的值和它们之间的连接强度计算总能量。

杰弗里·辛顿在霍普菲尔德网络的基础上创建了一种基于能量函数的概率模型,即玻尔兹曼机,它能够自主学习识别给定类型数据中的特征元素,在图像识别领域得到了广泛应用。除玻尔兹曼机外,1986 年,辛顿等人在《自然》上发表



诺贝尔科学奖解读

人工智能将成为推动各行各业变革的重要驱动力 (本图例由AI生成)



一个简单的神经网络(左)与一个包含多个隐藏层的深度神经网络(右)



诺贝尔物理学奖

▶ 本报特约作者 唐乾元

神经网络拥有独特模拟能力

拼图时尝试让每个拼块自己找到最合适的位置。这项研究不仅在理论上带来了对人神经网络如何存储和重构信息的全新视角,还为计算神经科学提供了重要的理论工具,帮助科学家探索大脑如何实现高效的计算和记忆。霍普菲尔德的工作还激发了后续关于关联记忆、模式识别和联想过程的深入研究,为理解大脑中的计算过程奠定了基础,也为后来出现的各种神经网络架构,类脑计算等提供了理论依据。

杰弗里·辛顿一直活跃于人工神经网络研究的第一线。从上个世纪 80 年代开始,他不断地在人工神经网络结构设计 and 训练方法上作出重要贡献,还推动深度学习的发展,或者说,深层次神经网络的复兴。2006 年,辛顿提出深度信念网络的概念,这是现代深度学习的基础。在一个更深层次的人工神经网络中,网络能够逐步提取到更为抽象和高层级的特征,从而从大量数据中学习到有意义的信息,这使得深度学习成为解决许多现实问题的关键工具。辛顿的工作帮助深度学习在计算机视觉和自然语言处理等领域取得突破性进展,为今天的人工智能奠定坚实的理论和

神经网络拥有独特模拟能力

20 世纪 80 年代初,当时人工智能的主流是基于符号逻辑的专家系统,例如,基于语法规则实现的机器翻译,这些系统依靠大量手工编写的规则和试图用逻辑推理模仿人类的思维过程,而神经网络的研究在那个时候并不受重视,很多人认为它缺乏实用性。约翰·霍普菲尔德是一位具有突破性进展,为今天的人工智能奠定坚实的理论和

灵感来自人类大脑结构的人工神经网络

▶ 本报特约作者 黄鑫

神经网络拥有独特模拟能力

论文《Experiments on Learning by Back Propagation》(反向传播学习实验)),让训练多层神经网络的“反向传播算法”广为人知;2006 年,辛顿在《科学》上首次提出“深度信念网络”概念,与传统的训练方式不同,“深度信念网络”有一个“预训练”(pre-training)阶段,这可方便地在神经网络中的权值找到一个接近最优解的值,之后再使用“微调”(fine-tuning)技术对整个网络进行优化训练。这两个技术的运用大幅度减少了训练多层神经网络的时间。他给多层神经网络相关的学习方法赋予了一个新名词——“深度学习”。辛顿在神经网络和机器学习方面的开创性工作,深刻影响了现代人工智能的发展,他被世人誉为“深度学习之父”。

辛顿是在获得图灵奖后又被授予诺贝尔奖。此前,仅赫伯特·亚历山大·西蒙(Herbert Alexander Simon,中文名司马贺)获得过此两项奖。在计算机科学方面,西蒙与艾伦·纽厄尔(Allen Newell)共同开发了“逻辑理论家”(Logic Theorist)和“通用问题求解器”(General Problem Solver)通用程序,开创了人工智能的研究方法。这些早期的人工智能系统不仅展示了计算机在问题求解中的潜力,也为后来的机器学习与智能系统的研究奠定了基础。1975 年,西蒙和艾

神经网络拥有独特模拟能力

使得人工智能能一夜间无人问津。近 20 年来,在摩尔定律的推动下,互联网、大数据和高性能计算机的发展为人工智能第三次产业浪潮奠定了坚实基础,语音识别和图像识别作为先锋队打响了人工智能回归大众视野的第一枪。2012 年,辛顿带领团队提出的 AlexNet 以深度学习预训练模型研究为代表的神经网络视觉识别研究迎来空前繁荣,众多落地应用普惠大众,具有较高认可度和良好前景。然而,历史上的人工智能研究并非一帆风顺,历经数次起伏。业界一些学者将人工智能的发展历程归纳为“三次浪潮,两次寒冬”,自从 1950 年艾伦·图灵(Allen Turing)提出“图灵测试”向世人抛出“机器能思考吗”的人工智能概念以来,这一领域产生了 1956-1974 年、1980-1990 年以及 2006 年至今的三次产业浪潮。前两次浪潮极大地推进了数学定理证明、问题求解、博弈、专家系统等方面的重大突破和进展,但每一次浪潮都险些被计算架构的神奇魔力,AlexNet 因此成为深度学习和人工智能自第二次“人工智能寒冬”后重要的里程碑,引发了深度学习革命。

神经网络拥有独特模拟能力

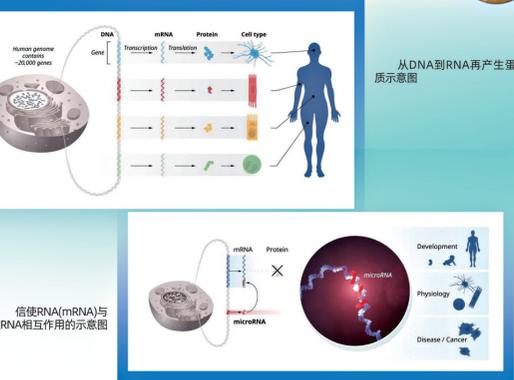
使得人工智能能一夜间无人问津。近 20 年来,在摩尔定律的推动下,互联网、大数据和高性能计算机的发展为人工智能第三次产业浪潮奠定了坚实基础,语音识别和图像识别作为先锋队打响了人工智能回归大众视野的第一枪。2012 年,辛顿带领团队提出的 AlexNet 以深度学习预训练模型研究为代表的神经网络视觉识别研究迎来空前繁荣,众多落地应用普惠大众,具有较高认可度和良好前景。然而,历史上的人工智能研究并非一帆风顺,历经数次起伏。业界一些学者将人工智能的发展历程归纳为“三次浪潮,两次寒冬”,自从 1950 年艾伦·图灵(Allen Turing)提出“图灵测试”向世人抛出“机器能思考吗”的人工智能概念以来,这一领域产生了 1956-1974 年、1980-1990 年以及 2006 年至今的三次产业浪潮。前两次浪潮极大地推进了数学定理证明、问题求解、博弈、专家系统等方面的重大突破和进展,但每一次浪潮都险些被计算架构的神奇魔力,AlexNet 因此成为深度学习和人工智能自第二次“人工智能寒冬”后重要的里程碑,引发了深度学习革命。

神经网络拥有独特模拟能力

▶ 本报特约作者 唐乾元

神经网络拥有独特模拟能力

诺贝尔生理学或医学奖



神经网络拥有独特模拟能力

使得科学家们能够处理大规模的数据,结合科学理论进行更精确的预测和效率。这种跨学科的神经网络在图像和语音识别任务中表现出色,让全球范围内的科学家们重新认识到神经网络的潜力,也点燃了科学界的广泛兴趣。

从技术应用的视角看,深度学习已成为推动人工智能进步的核心技术,广泛应用于未来产业和日常生活中的各种场景。例如,图像和语音识别、智能手机中的语音助手、自动驾驶技术、图像与视频生成技术等都得益于深度学习的突破。

深度学习的应用改变了科技产业,还推动了多个传统行业的变革。例如,在医疗行业中,深度学习用于医学图像分析,帮助医生进行更精确的诊断,从而提高治疗的效率和准确性;在制造业中,深度学习被用来实现生产过程的自动化和智能化,提升生产效率和产品质量;在能源行业中,深度学习被用于预测电力需求和优化能源分配,帮助提高能源使用效率,减少浪费;在零售行业,深度学习被用于顾客行为分析和个性化推荐,帮助商家更好地理解顾客需求,提供更加精准的服务;在物流行业,深度学习被用于优化运输路线和仓储管理,通过实时分析交通情况和库存数据,提高配送的效率和降低成本。

上述这些应用使人工智能成为推动各行各业变革的重要驱动力。这些突破将改变人们看待世界的方式,推动科技进步,为人类社会带来深远的意义。

神经网络拥有独特模拟能力

(作者:唐乾元,博士,香港浸会大学物理系助理教授)

神经网络拥有独特模拟能力

人与机器的和谐共生。这一奖项的颁发传递了一个强烈的信号:人工智能不再是科幻小说中的人类幻想,它正在逐步融入人类社会的每一个角落,成为推动社会进步的重要力量,促使社会各界开始更加深入地思考人工智能的伦理、法律和社会影响,增强跨学科、跨国界的合作与交流,共同构建一个更加包容、公正、可持续的智能社会框架。在这一过程中,约翰·霍普菲尔德和本弗里·辛顿的成就就鼓励更多科学家、工程师和社会学家投身于人工智能的持续健康发展之中。

未来高新技术产业将受益于人工智能技术的不断成熟与普及,人工智能将在与其他产业的深度融合后,将催生出更多新兴业态和商业模式,譬如为交通运输等传统行业带来革命性变化。在雄安新区,无人驾驶巴士已成为常态,其反应速度远超人类驾驶员,凭借高科技传感器实现了车辆周边 240 米范围的环境感知。通过多功能信息杆和智慧运营中心系统,雄安新区还构建了智慧路灯、智慧斑马线等智能交通设施。这些设施利用大数据智能分析和多传感器融合技术,实时采集道路信息,为交通信号协同控制提供了重要数据支撑。

约翰·霍普菲尔德与本弗里·辛顿荣获 2024 年诺贝尔物理学奖,是对人工智能领域发展成就的肯定,也是对未来无限可能的期许。它标志着人类社会正站在一个崭新的起点上,准备迎接由人工智能引领的科技革命,共同创造一个更加智慧、高效和和谐的世界。

(作者:黄鑫,博士,交通运输部规划研究院工程师)

神经网络拥有独特模拟能力

本版主编 策划 朱广清

神经网络拥有独特模拟能力

神经网络拥有独特模拟能力