

神经网络大模型: 破解生命密码的新钥匙

□ 王琳琳

近日,谷歌深脑团队(DeepMind)和同构公司(Isomorphic)在国际科技期刊《自然》发表一项重大研究成果,揭示在蛋白质结构预测领域取得的突破性进展。科学家通过联合自注意力机制和扩散模块,推出一种新模型,可成功预测几乎所有生命分子的结构及其相互作用。这一突破让我们对生命中复杂现象的理解进入了一个全新的阶段。

AI大模型让药物研发更精准

相较于此前的蛋白质结构和相互作用预测工具,这项研究成果将成功率提升50%以上。它不仅可取代现有的许多结构解析工作,还能广泛应用于生物医药行业,大大加速了新型药物的开发进程,并显著降低开发成本。举个例子来说,在抗肿瘤免疫调节过程中,通过对相关蛋白结构预测,以及相应蛋白与新型小分子药物结合后的结构变化预测,可有效评估小分子药物的疗效潜力。这一过程仅需在服务器上输入蛋白质的基因序列和小分子结构式,即可快速完成预测。相比之下,传统方法则需进行大量耗时且昂贵的生物化学实验才能获得类似结果。

然而,生物大分子的结构多变与复杂相互作用,只是生命复杂性的冰山一角。为了真正理解生命如何从一个个单独分子形成细胞,进而形成个体并表现出生长发育、自我繁殖,以及适应环境等多种复杂行为,我们不仅需要掌握生物分子的结构和相互作用,还必须了解这些生命分子如何在生命体内的集体运动机制。

为了精确描述生命体中无数分子的集体运动机制,科学家们建立了数学模型来描述系统状态随时间的变化。这些模型被称为动力学模型。事实上,

在现实世界中,细胞中的分子会受到溶液、温度和其他分子的相互作用等因素干扰,导致我们无法精确预测这些分子的具体状态。为了应对这种不确定性,科学家们开始采用概率的方法,描述这种概率分布变化的模型,被称为随机动力学模型。

通过求解随机动力学方程,我们可以了解关于分子运动的所有信息。但是,概率性导致数据信息量变得巨大,需要计算的空间点的个数也随着分子数量的增加而呈指数级增加。例如,对100种化合物分子而言,就需要计算超过 10^{30} 个空间位置,这甚至远远超过了整个宇宙中的原子个数,得到所有空间位置概率所需的计算时间也与宇宙的年龄相当。因此,为了处理如此庞大的计算量,科学家们一直在探索更加高效的计算方法和模型,希望能更精准、更快速地揭示生命分子的动态行为。

AI大模型实现随机动力系统高效解算

为了破解这一计算难题,中国科学院长春应用化学研究所博士刘传波和国科温州研究院研究员汪劲在国际期刊《美国科学院院刊》上发表了一篇突破性的文章。他们提出了一种基于AI大模型的随机动力系统求解方法。通过使用基于强化学习的知识蒸馏技术,他们首次证明,可以在极高精度下将高维随机系统的全部动力学信息,编码到一个统一的神经网络中。这可以形象地比喻为将一个如同宇宙般庞大的空间压缩进一个神经网络大模型。此求解方法将原本指数级的计算复杂度压缩到近似线性,使得对包含超过100种化合物的随机动力系统的精确求解成



神经网络大模型如同黑洞一样提取并压缩随机动力系统信息。刘传波 制图

为可能,成功攻克了随机动力系统的计算难题。同时,他们还发现经过训练的神经网络获得了与大语言模型类似的泛化能力,可根据已学习过的方程精确预测其他动力学方程的概率。这可以类比于人类举一反三的能力,例如,学习了 $1+1=2$ 之后,就知道 $1+2=3$ 。而且神经网络比人类更加精确,展现出了超越人类的对随机动力系统的理解能力。通过应用这一新方法,未来科学家将能以前所未有的精度对整个细胞乃至整个组织进行定量分析,从而探索生命行为如何从一堆无生命的分子中涌现出来。

古希腊德尔斐神庙墙上镌刻着“认识你自己”的箴言。几千年来,这句古老的智慧名言激发了无数哲学家的深思与探讨。今天,随着人工智能大模型技术的飞速发展,这句话正获得全新的意义。借助AI大模型对生命系统中分子结构和相互作用的预测,以及对分子在细胞中运动的精确计算,人类终于可以逐步揭开生命的神秘面纱,逐渐认识真实的自己。

(作者系吉林省科普创作协会会员、吉林省科学技术工作者服务中心助理研究员)

凡尔纳“没有发明任何东西”吗

□ 尹传红



上月末,经过164天、总航程约5.7万公里的旅程后,“深海一号”船携蛟龙号载人潜水器回到位于青岛的国家深海基地码头,圆满完成了我国首个大西洋载人深潜科考航次任务。

此次考察开展载人深潜作业46次,创造了单航次下潜次数最多的纪录,获取了热液喷口区域温度、甲烷和氢气等第一手环境数据,采集到盲虾、贻贝和蠕虫等宝贵的深海热液生物样品以及硫化物、玄武岩和沉积物等深海地质样品,对我们未来的深海科学研究意义重大。

读到媒体相关报道之后没多久,我在北京理工大学出版社科幻分社,看到了即将付印的《海底两万里》中译本精美的科普图解版,真是浮想联翩。距我第一次如痴如醉地阅读儒勒·凡尔纳这部科幻经典,已然流逝了40多年光阴,仿佛自己也随着科幻走进了现实。

凡尔纳著名的科幻三部曲《格兰特船长的女儿》《海底两万里》和《神秘岛》对地球和海洋探索具有史诗般寓意,产生了世界性的广泛影响。三部曲中,《海底两万

里》的设想或许是最早成型。大约在1866年,凡尔纳“想到了一个好主意”,构思出一个陌生人“必须跟他离开的人类不发生任何关系。他不再到陆地上来,他不把陆地放在眼里。他只需要大海,而大海必须能够向他提供一切……”

客观地讲,早在凡尔纳出生之前,一些能潜入水底的器械就已经发明出来了。只是,由于技术上的原因,这类“海底运动器”一直未能产生实际的应用价值。借助《海底两万里》这部科幻小说,凡尔纳展开了他大胆、神奇的想象,详细地描绘了他“创造”的鹦鹉螺号潜水艇的构造和各种功能,以及人们经由它而得以窥见的丰富无比的海底世界。

事实上,在《海底两万里》问世30年后,世界上第一艘具有现代意义的潜水艇才被发明出来。美国科学家、潜艇发明者西蒙·莱克在其自传中开篇即写道:“儒勒·凡尔纳是我一生事业的总指导。”而1958年首次抵达北极的原子能潜水艇,就是以鹦鹉螺号命名的。

不过,1903年,75岁的凡尔纳在接受一位英国作家采访时谈到:“我的主人公尼摩是一个愤世者,希望跟陆地一刀两断,他从海洋中获得他的动力:电力。这是有科学依据的,因为海洋包含电力储备,就像地球一样,但获取这种

力量的方法从未被发现。所以我没有发明任何东西。”

后人评价凡尔纳这段自谦之语——“我没有发明任何东西”时,认为凡尔纳的创作其实具备强有力的现实依托,小说中使用的素材大多来自既有的科学成果,并在此基础上进行了合理的推演,而非无中生有的“发明”。凡尔纳本人在那次受访后一年,即他离世(1905年)的前一年,更为明确地指出:

以鹦鹉螺号为例,如果认真想一想的话,这就是一种无甚特别之处的水下机械,也没有超出当前科学知识的边界。它通过完全可行而且众所周知的过程上浮或下潜,其制导和推进方面的各种细节全都完全合情合理而且可以理解。它的动力甚至都不是秘密:我求助于想象力的唯一一点是这种力量到底如何应用,在这方面我故意进行了留白,让读者形成属于他自己的结论,这只是一个技术方面的空隙,完全可以由训练有素、经验丰富的头脑加以填补。

由《海底两万里》这个事例可以看到,一个多世纪以来,科幻作品和发明创造一直存在着千丝万缕的密切联系。某些技术在实验室产生的同时,也出现在幻想作品中。苏联有个叫作根里希·阿奇舒勒的发明家和科幻作家,

潜心研究了20世纪中期数千项专利,试图发现与人类想象力有关的线索,并提出了一些方法,用来预测未来科技走向,解决技术难题。他坦陈,是凡尔纳的作品激发了自己的想象力。

14岁时,阿奇舒勒受《海底两万里》中所描述技术的鼓舞,开始为自己的发明制订计划。他想到,尼摩船长能够在海底行走,所以我们需要有耐压服,于是他就设计了一套潜水衣。在后来的深入研究中,阿奇舒勒意识到,实际上是凡尔纳首次提出了潜水艇双层船体的概念,而且其小说所涉及的细节内容完全体现在后来的专利发明中,可凡尔纳并不是相关专利权持有人。为此,阿奇舒勒开始倡导“科幻思想登记制度”——这有些类似于针对幻想技术的专利制度。

阿奇舒勒或许称得上是制定发明规则的第一人。该规则涉及科幻式的预测、认知科学,以及对于技术体系发展史的深刻理解。他的发明方法和理论,后来在英语国家被称为“创新问题解决理论”(TRIZ)。他认为发明活动可以后天习得,可以加以改善、提速和改进。他相信:具备创新能力的人越多,社会就会越美好、越高级。

由阿奇舒勒的事例看,凡尔纳“没有发明任何东西”吗?