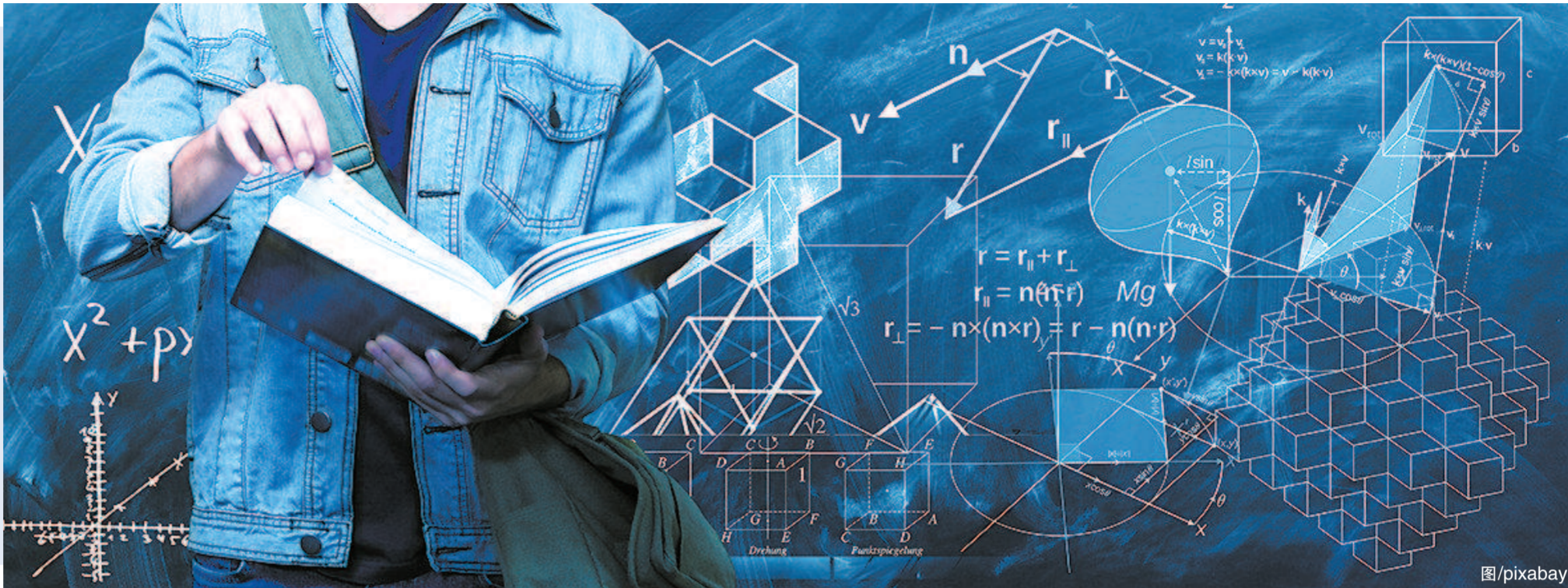


科学家正努力让AI背后的深度学习算法既能回答“是什么”又能说出“为什么”

发现自然界新定律，“机器理论家”呼之欲出



图/pixabay

人工智能(AI)正在推动科研范式变革,快速且深刻地影响着物理学、化学、材料学、生物学等领域。随着人类探测宇宙本领的提升,高灵敏、高精度的望远镜、卫星不断产生着海量数据,而这正是AI大展身手的天地。

目前,AI算法已可通过天体物理学的数据集来确定一些基本方程,它能就此发现新的物理学定律吗?事实上,这正是物理学家努力的方向,他们正在研究如何才能让这些“机器理论家”具备发现自然界更深层次规律的能力。

——编者

■林宇/编译

1980年,著名物理学家斯蒂芬·霍金在英国剑桥大学发表演讲时,提出了“万物理论”,将科学家对客观世界的两种主要描述方式——广义相对论和量子力学,归结为一个简洁的、包罗万象的方程式,并对未来人工智能不断增长的能力作出了预测。

自霍金提出“万物理论”以来,人工智能取得了飞跃式发展,但物理学家在利用机器学习来寻找新的、更深层次的自然规律方面,却仍然进展缓慢。AI背后的深度学习算法通常只能回答“是什么”,

却无法说出“为什么”。

如今,AI已成为科学家从海量数据中寻找和发现新方程式的得力助手。美国西雅图华盛顿大学的史蒂夫·布伦顿说:“我们正在进入这一发现阶段。”而在发现新的物理学定律方面,AI将成为理论物理学家的最终合作者。现在,物理学家面临的挑战是,如何让机器避开人类理论物理学家的固有思维,从而找到人类找不到的定律。

为此,物理学家和人工智能正在寻找某种共同语言,一种真正能够实现人机对话的新语言,从而在人类与AI之间形成真正的合作关系。这将是一场重大变革的开始。

算法优势互补,“重新发现”万有引力定律

数据驱动的科学探索始于16世纪。当时,荷兰天文学家、近代天文学奠基人第谷·布拉赫对行星和恒星的运动进行了细致观察。到了17世纪,德国天文学家、数学家和物理学家约翰内斯·开普勒仔细研究了布拉赫的观察笔记,据此提出了三个简洁的定律来描述行星的运动,确立了太阳为太阳系中心的理论。

在研究这些数据时,开普勒用的是试错法。他煞费苦心研究了无数的轨道形状,以寻找最适合太阳系的轨道形状。最终,他发现了行星运行轨迹与绕太阳公转所需时间之间的精确数学等式,首次将数学方程式运用于人类对宇宙的理解。

与开普勒同时代的太阳中心论者伽利略曾说:“描述自然的最伟大的著作是用数学语言写成的。”可是,数学定律真的本就是宇宙运行的固有特性?抑或是我们强加于它的?

这个问题至今未有定论。考虑到数学在用简单方程式捕捉抽象定律方面所取得的惊人成就,现代物理学家想尝试让人工智能来做类似的事情,并非没有道理。

基于这样的设想,美国卡内基梅隆大学的帕特里克·兰利于上世纪70年代首次提出“符号回归”(Symbolic regression)算法。符号回归是一种机器学习算法,通过搜索并识别潜在的数学表达式,来寻找最适用于给定数据集的模型。

符号回归的工作原理是系统地运行带有各种数学符号或运算符的方程式,包括加减法和各种物理变量(位置或速度等)的组合。如果其中某个方程式与数据非常吻合,如行星轨道的观察数据,人工智能就会得到奖励。然后,算法将通过改变正负号等方法产生新的方程式,再对数据集进行测试,并与之前的方程式比较,逻辑不周密或缺乏依据的方程式会逐渐被淘汰——这个过程与自然选择相类似。

符号回归必须在高维数据集内寻找新的方程式。这些数据集的一大特点就是包含许多物理变量,比如广受天体物理学家和宇宙学家青睐的来自韦伯空间望远镜、欧洲航天局盖亚太空天文台等大望远镜的大量数据。

对于符号回归算法来说,变量多是一个棘手问题——太多变量导致需要测试的潜在方程式的数量呈爆炸式增长,即便对目前性能最强的超级计算机来说,这也会是一个难以承担的任务。

但是,深度学习算法的优势就是可以轻松处理庞大的数据集。为此,早在几年



▲通过对宇宙的高精度观测,现代大型望远镜和卫星正产生大量数据。物理学家希望通过AI,从中发现更多自然界的新定律。
▶詹姆斯·韦伯空间望远镜传回海量宇宙观测数据。图/视觉中国

前,美国普林斯顿大学的迈克尔·克兰默,就与美国纽约喷气推进实验室(专注于计算科学的研究)的雪莉·何合作,尝试结合两种方法的优势,以充分利用深度学习的模式发现能力,以及符号回归于解释和理解的输出结果。

同时,符号回归也是对深度学习模型输出的一种自然补充。深度学习模型的输出往往难以解释,而符号回归作为数学语言中更易于理解的输出,可以帮助物理学家产生新的可验证假设。

两位科学家将美国航空航天局(NASA)关于太阳系行星及其卫星运行轨道的数据(当年开普勒计算研究的内容),以及一系列数字和符号喂入深度学习神经网络。如果神经网络发现数据中存在某种模式,就会将其确定为潜在目标,最终输出的结果将形成一套新的数据集。

正如研究人员所料,简称为“PySR”的符号回归算法通过这些原始数据,果然“重新发现”了牛顿的万有引力定律。

从原则上证明了这一方法的可行性后,克兰默和其他一些研究人员利用大量新的天体物理学数据,通过PySR来发现和描述宇宙多样性及相互关联特征的方程式。

从引力波探测推演黑洞质量,到描述宇宙空洞特征,符号回归算法为天体物理学家提供了寻找宇宙数学秩序的新方法。“靠数据来推动新发现,这真的很厉害。”布伦顿说,他在2016年与人合作创建了另一种流行的符号回归算法SINDy。

数学符号与数据拟合并不是我们理解世界的唯一方法。爱因斯坦提出广义相对论,就是通过一系列富有想象力的思想实验而实现的



图/pixabay

未来十年内,AI或将带来重大物理学发现

在宇宙中占比80%的暗物质是一种神秘的引力来源,其引力可维持星系不会飞离。在宇宙学中,暗物质缺失的地方被称为宇宙空洞,空洞的分布和特征可能与宇宙常数有关。但要发现相关的方程式很难,因为宇宙中有很多空洞,每个空洞的描述都不一样,所以有很多变量。

2021年5月,克兰默与雪莉·何等其他合作者一起,结合深度学习神经网络与PySR符号回归算法,发现了宇宙空洞大小和形状与宇宙总能量之间的比例关系。

2022年3月,美国普林斯顿大学的海伦·绍和她的合作者利用PySR符号回归算法,发现了从恒星形成速度预测亚星质量的方程式(亚星是指形成密集团块的暗物质)。令人惊讶的是,这个方程式能够准确适用于宇宙中几乎所有的星系。

这些具有普遍性的宇宙基本关系,正是物理学家努力寻找的。它们可以用来描述很多不同寻常的物理系统,也是人类理解宇宙特征的标尺。正如牛顿第二运动定律(即作用于物体上的力等于物体的质量乘以加速度)既适用于牛顿最初发现的苹果落地地奥秘,也适用于人类发射火箭、着陆月球等壮举。

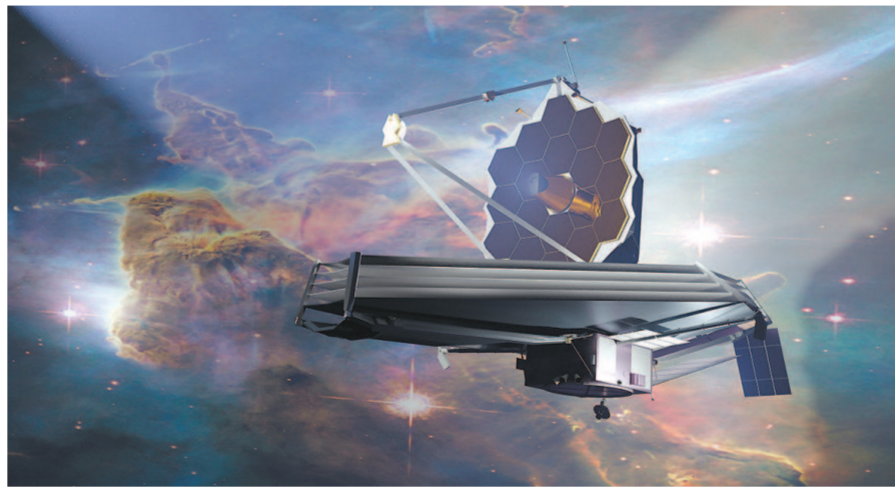
符号回归算法往往可以发现更具普遍

性的方程,而深度学习神经网络一旦遇到超出其理解范围“舒适区”的情况,往往会输出一些毫无意义的结论。

但到目前为止,PySR发现的都是描述性的经验方程式——擅长于复制实验数据,而不能直接提供物理学家想要的理论解释,或解答更深层次的“为什么”。例如,开普勒定律就是一个经验方程式,他的方程式与布拉赫观测到的大量数据惊人吻合,但开普勒的方程式无法回答“为什么”会吻合。

我们知道,数学符号与数据拟合并不是我们理解世界的唯一方法。爱因斯坦提出广义相对论,就是通过一系列富有想象力的思想实验而实现的。而一些观测结果,比如水星在夜空中不寻常的运动,只是证实了广义相对论中“引力是质量扭曲时空的结果”的论断,而不是激发了这一理论的产生。

克兰默认为,经验方程式是通向更深层次定律的跳板,而非宇宙真相本身。他甚至不愿称这些新方程式为“发现”,“我们肯定还没有到那个地步”。据他推测,符号回归算法和人工智能深度学习可能在未来十年内带来重大科学发现,比如暗物质究竟是什么,或者它是否真的存在。



避免思维固化,有助找到人类找不到的方程

符号回归算法可为物理学家指引方向,帮助他们实现更大的飞跃。但在研究这些方程式——理解它们的形式、彼此之间的关系,或整个物理学之间的关联方面,科学家仍发挥着至关重要的作用,至少目前如此。

加拿大蒙特利尔大学的帕布罗·莱莫斯也参与了重新发现牛顿万有引力定律的项目,“我们现在显然遇到了瓶颈,符号回归也许能够给我们带来新的启示。”

对于如何开发能够自主发现基本自然规律的人工智能,克兰默等研究人员有了新的设想。

符号回归算法通常要在准确性和简单性之间做出权衡。如果方程式非常复杂,与所用数据集匹配度很高,通常被称为“过度拟合”,会增加数学表达式在测试数据集外不准确的危险,换句话说,不具备普遍性。而简单表达式则“更有可能真正捕捉到某种机制”。故而,在PySR和其他符号回归程序中,一直流行一种简单化倾向,即去除不必要的复杂解释,以及去除方程式中的冗余符号。当时,PySR利用NASA的轨道数据集重新发现的牛顿的万有引力定律,而不是爱因斯坦的广义相对论方程式,可能正是出于这样的权衡。

对称性是物理学家寻找普遍自然法则的另一盏指路明灯。布伦顿和美国华盛顿大学的南希·库兹一起,将多种类型的对称性知识“嵌入”SINDy算法程序中,用来从描述流动流体复杂行为的数据中提取方程式。“你几乎总是能从更少的数据中得到准确性更强的模型,将更多的垃圾信息剔除出去。”布伦顿说。

“数学之美”虽然难以定义,但它也是许多物理学家努力追求的一个目标。如

今,物理学家的这些想法都纳入了人工智能寻找新的数学方程式的范畴。

美国麻省理工学院的杰西·塞勒指出,尽管我们对现实世界的直觉在过去曾经发挥过重要作用,但直觉有时也会让我们步入歧途,甚至迷失方向。

为此塞勒警告说,如果试图让人工智能去发现,“我们可能会错过它们最具革命性的前景——提供新视角的能力”。莱莫斯也认同这一观点,“我们不想给它太多类似于人类的直觉,因为我们的最终目标是利用它来找到人类找不到的方程。人类的直觉也许会让它像我们一样遭遇同样的瓶颈”。

塞勒记得,有一次,他让人工智能解决困扰了他十年之久的一个难题。不出他所料,AI很快就反馈出解决方案。经过仔细核对检查后,他为自己之前为何没有想到而感到尴尬,“我终于明白,是人类固有的思维方式限制了自己的思路。所幸的是,我没有将自己的这种思维方式加之于人工智能。”

AI需要在自由王国里发挥它们的强大能力,通过测试数学符号的每一种不同组合,搜索解决问题的所有可能性。塞勒认为,物理学家也需要转换角度去思考,“如果我们想让人工智能在基础物理学上取得重大飞跃,这正说明我们自己的头脑需要转变”。

目前,符号回归正在赢得人工智能怀疑论者的支持。从某种意义上说,符号回归只是望远镜、计算机和微积分之外的另一工具。计算机并非天生好奇,因此需要人类来向它提问。“人类仍然必须为它指定操作的沙箱,然后让计算机为我们查看并检查沙箱中的每一粒沙子。”塞勒说。

■宇辰/编译

只要描述你想要看的东西,就可以制作出几乎以假乱真的高质量视频,无论是基于现实还是完全的异想天开。将文本转换成视频的人工智能算法一经问世,就立刻蹿红。而我们可能由此进入了一个“眼见不再为实”的世界。这项技术的诞生,无疑会对创意产业产生深远影响,但同时也可能导致网络媒体的信任危机。

有限学习资源限制技术完善

目前,Meta公司的Make-A-Video、谷歌的Imagen Video所生成的视频质量还较差,在努力复制真实世界时,有不少会留下人为伪造的怪异视觉印象。

但早期的文本到图像模型同样也非常粗糙,非常不真实,不过图像质量很快得到了改善。眼下,从文本到视频模型需要解决的一个关键问题是,对图像运动和时流流逝关系的更深理解。

“你得到的是即时快照,不包含接下来会发生什么的信息,模型系统不理解因果关系,也不懂物理学常识。”Stability AI首席技术官汤姆·梅森表示。

另外,从文本到视频模型缺乏详细的、有标签的数据集,而文本到图像模型可以依赖于庞大的数据集来进行训练。Make-A-Video只能让人工智能以“无监督”学习的方式,在大型视频网站上浏览扫描。Meta AI研究总监德维·帕里克指出,人工智能从无监督学习模式中得到的东西可能非常有限。他认为,文本到视频模型目前尚未能达到文本到图像模型的水平。

“眼见不为实”带来风险隐忧

即便如此,对于新出现的从文本到视频技术,人们已经开始产生了一些担忧。用来让人工智能学习的视频中,如果存在任何偏见,都有可能通过模型传播开来,并出现在后续生成的视频中。

“你会希望你的孩子仅仅通过网上视频,就能学到他们所需知道的人类社会的一切吗?”卡内基伦理与国际事务委员会的亚瑟·霍兰德·米歇尔问道。

在错误信息和误导性信息泛滥的网络世界里,偏见并非唯一的隐忧。人工智能创造完美视频的能力越强,就越能以假乱真,让人难辨真假。

首次人工智能驱动的视频篡改事件,早在几年前就曾出现。例如,深度伪造(deepfakes)技术可以把一个人的头部照片安放在另一个人的身体上。当时就有人提出警告:这种技术将导致网络媒体信任度下降。

过去,要制造以假乱真的效果并非易事,因为那时还需要高超的技术能力。但如今有了从文本到视频生成模型这一利器,技术上的易懂性和易用性将会大大提高。换言之,只要是稍微懂一点这方面技术的人,就有可能创建出可信度极高的造假视频。

为应对深度造假和人们对文本转视频的担忧,英国斯旺西大学的伊冯·麦克德蒙特·里斯和她的同事们启动了一项为期五年的项目,对人们网络媒体态度的变化进行评测。研究表明,人们在网上传播和分享哪种媒体往往与个人的先入之见有关,哪怕看出视频可能造假。

立法能否跟上技术进步

当被问及Make-A-Video是否会被用来传播虚假信息时,Meta的一位发言人表示:“作为研究的一部分,我们将继续探索,进一步完善和降低潜在风险的方法。”例如,对AI培训数据进行严格审查和过滤,以减少它们接触到有争议、有问题的内容。

除了可能促成虚假信息泛滥,从文本到视频的技术还将对创意产业造成极大影响,因为许多图像和视频都是版权作品。从研究成果转向商业产品,人工智能的输出结果也将面临法律上的挑战。虽然目前为止还远没有出现相关的诉讼案例,但训练人工智能或人工智能输出的过程,都有可能成为一种新形式的版权侵权。

这意味着可能需要引入新的相关法律,然而目前尚不清楚立法是否能跟上人工智能技术的发展速度。英国纽卡斯尔大学的莉莲·爱德华兹说:“在某种程度上,法律已经做了能做的大部分事情。作为一种现象,假新闻普遍存在的一个问题是,它‘并不明显违法’,但却带来危害。”她说,从文本到视频的人工智能输出也存在同样的问题。为了保持人们对网络媒体的信任度,一种可能的解决方法是使用认证证书,以明确视频的来源和创建方式。

随着文本生成视频技术的商业化,新型侵权及真实性担忧随之而来 视频不再『眼见为实』,信任度如何维持