

寻找新粒子、分析公众情绪、梳理病因……

看人工智能如何改变科研

寻找新粒子

上世纪80年代末,正当“神经网络”这一术语捕获了公众的想象力时,粒子物理学家开始“玩弄”起人工智能。他们的领域很适合应用人工智能和机器学习算法,因为几乎每项试验均聚焦于从复杂粒子探测器获取的不计其数的类似数据中寻找微妙的空间模式,而这正是人工智能擅长的事情。“我们花了好几年时间才说服人们,人工智能并不是一种神秘的把戏。”最早拥抱此项技术的物理学家之一、美国费米国家加速器实验室研究人员 Boaz Klima 表示。

粒子物理学家力图使带有巨大能量的亚原子粒子相撞以迸发出独特的新物质微粒,从而理解宇宙的内部运作。例如,2012年,利用全球最大质子对撞机——位于瑞士的大型强子对撞机(LHC)开展研究的团队发现了预期已久的希格斯玻色子。这是一种转瞬即逝的粒子,对于物理学家解释所有其他基本粒子如何获得质量至关重要。

不过,此类奇特粒子并非自带标签。在LHC,几乎约10亿次对撞才出现1个希格斯玻色子。与此同时,它会在十亿分之一皮秒内衰变成诸如光子等其他粒子。为“重现”希格斯玻色子,物理学家必须认出所有那些更加常见的粒子。然而,典型碰撞中产生的成群的不相关粒子使这项工作变得更加艰难。

费米实验室物理学家 Pushpalatha Bhat 介绍说,诸如神经网络等算法擅长从背景中筛选信号。在粒子探测器(一般是由各种传感器构成的巨型筒状集合体)中,光子通常在被称为电磁量能器的子系统中创建粒子束。电子和强子也是这样产生的,但它们的束流和光子稍微有些不同。机器学习算法通过发现描述束流的多个变量之间的相关性,将它们区别开来。此类算法还能帮助区分希格斯玻色子衰变产生的光子对。“这是一个公认的大海捞针式的问题。”Bhat表示,“这也是为何我们要从数据中提取尽可能多的信息。”

然而,机器学习并未完全占领这个领域。物理学家仍然主要依靠对基础物理的理解来断定如何搜索数据,以寻找新粒子和现象存在的迹象。不过,劳伦斯伯克利国家实验室计算机专家 Paolo Calafiura 表示,人工智能可能正变得愈发重要。到2024年,研究人员计划升级LHC,从而使其碰撞率提高10倍。Calafiura 介绍说,到那时,机器学习将在应对数据洪流时发挥至关重要的作用。

分析公众情绪

伴随着每年几十亿用户以及数千亿条推特和帖子的产生,社交媒体已将大数据带入社会科学。同时,心理学家 Martin Seligman 表示,它还为利用人工智能收集人类传播产生的意义创造了史无前例的机遇。在宾夕法尼亚大学正向心理学中心,Seligman 同来自“全球福祉项目”的20多名心理学家、内科医生和计算机专家,利用机器学习和自然语言处理筛选大量数据,以估量公众的情感和身体健康。

传统上,这是通过调查实现的。不过,



人工智能工具正在帮助揭示可能影响自闭症的上千个基因。

图片来源:BSIP SA/ALAMY STOCK PHOTO

Seligman 表示,社交媒体数据“比较低调”、花费较少,而且获得的数据要高好几个数量级。虽然此类数据也很散乱,但人工智能提供了一种强有力的获取模式。

在最新研究中,Seligman 和同事分析了2.9万名自我评估患有抑郁症的脸书用户更新的内容。利用来自其中2.8万名用户的数据,机器学习算法发现了更新内容中的词语和抑郁症水平之间的关联。随后,它仅基于更新的内容,成功估量出其他用户的抑郁水平。

在另一项研究中,该团队通过分析1.48亿条推特,预测了县级心脏病死亡率。事实证明,同愤怒和负面情绪相关的词语是危险因素。和基于诸如吸烟、糖尿病等10项主要危险因素的预测相比,这项来自社交媒体的预测同实际死亡率匹配得更加紧密。与此同时,研究人员利用社交媒体预测了个性、收入和政治意识形态,并且研究了住院治疗、神秘体验和刻板印象。该团队甚至利用从推特上推断出的福利、抑郁症、信任和五大人格特征,创建了一幅为美国每个县作出标识的地图。

“在分析语言及其同心理学的联系方面,一场革命正在上演。”得克萨斯大学社会心理学家 James Pennebaker 表示。Pennebaker 的研究并非内容,而是风格。他发现,在申请大学入学短文中虚词的使用能预测成绩。冠词和

介词象征着分析性思维并且预示了更好的成绩,代词和副词象征着叙事性思维并且预示了较差的成绩。Pennebaker 还发现了证据,表明1728年的剧本《双重背叛》的大部分内容可能由莎士比亚撰写。机器学习算法基于诸如认知复杂性和罕见词等因素,将该剧本同莎士比亚的其他作品进行了匹配。“现在,我们可以分析你贴出甚至撰写的任何内容。”Pennebaker 表示,结果是“获得了关于人们是什么样子的愈发丰富的画面”。

探寻自闭症根源

对于遗传学家来说,自闭症是一项棘手的挑战。遗传定律表明,它拥有强大的遗传因素。然而,已知在自闭症中起到一定作用的许多基因的变体只能解释约20%的病例。寻找可能影响自闭症的其他变体,需要从关于2.5万个其他人类基因及其周围DNA的数据中搜寻线索。对于人类研究者来说,这是一项艰巨的任务。为此,普林斯顿大学计算生物学家 Olga Troyanskaya 和纽约西蒙斯基金会取得了人工智能工具的支持。

Troyanskaya 将关于哪些基因在特定人类细胞中活跃、蛋白如何相互作用以及转录因子结合位点和其他关键基因组特征位于哪里

的上百个数据集结合在一起。随后,她的团队利用机器学习构建了基因相互作用的地图,并且将已得到确认的较少自闭症危险基因同上百个涉及自闭症的其他未知基因进行了对比,以寻找它们的相似性。此项研究标记了另外2500个可能同自闭症相关的基因。相关成果在去年发表于《自然—神经科学》杂志。

不过,正如遗传学家最近意识到的,基因并不是孤立地发挥作用。它们的行为受到上百万个附近非编码碱基的影响。这些非编码碱基同DNA结合蛋白以及其他因素相互作用。确认哪些非编码变体可能影响附近的自闭症基因是一个更加棘手的问题。Troyanskaya 的研究生 Jian Zhou 正在利用人工智能解决这一难题。

为训练程序(一个深度学习系统),Zhou 将其暴露在“DNA 元件百科全书”和“表观基因组学路线图”收集的数据中。这两个项目对上万个非编码DNA位点如何影响附近基因进行了梳理。Zhou 利用的系统学习了在评估未知非编码DNA的潜在活性时应该寻找哪些特征。

当 Zhou 和 Troyanskaya 于2015年10月在《自然—方法学》杂志上描述了这个名为 DeepSEA 的项目时,加州大学尔湾分校计算机专家 Xiaohui Xie 将其称为“将深度学习应用于基因组学的里程碑”。目前,该研究团队正通过 DeepSEA 研究自闭症患者父母的基因组,以期对非编码碱基的影响进行排序。(宗华编译)

臭氧层新杀手浮出水面

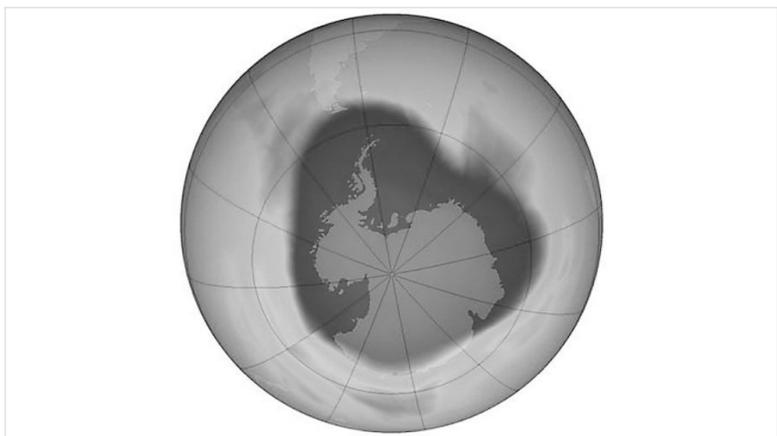
或使南极臭氧层恢复延迟5到30年

臭氧层(保护人体免受太阳紫外线辐射的高海拔氧分子层)在过去10年间一直在恢复。但一种新发现的威胁可能延迟其恢复。研究人员发现,过去几年,溶剂、油漆去除剂和药品生产中经常使用的一种化学物质的排放已达到原来的两倍,这将会导致南极洲的臭氧层恢复延迟5至30年,如果该化学物质水平继续增长,未来恢复期甚至可能更长。

对此,澳大利亚墨尔本大学环境科学家 Robyn Schofield 说,这些发现“令人觉得可怕”,这是一件大事”。

这种受质疑的物质是二氯甲烷。英国兰卡斯特大学大气化学家 Ryan Hossaini 说,这种物质的自然界来源极少,因此近年来其排量的增加可能来自于人类。他补充说,从2000年到2012年,平均来看二氯甲烷浓度增加了。在全球范围内,从2004年到2014年,二氯甲烷浓度增长接近原来的两倍。Hossaini 及其团队推测,目前二氯甲烷排放为每年约100万吨。

像含氯氟烃和其他破坏臭氧层化学物质一样,二氯甲烷在遇到阳光时会分解。氯原子被释放后会消除任何与其发生相互作用臭氧分子。1987年,《蒙特利尔议定书》达成对生产和使用含氯氟烃以及许多与工业化国家相关的化合物的禁令,但却忽视了二氯甲烷,因为研究人员认为它不能原封不动地在大气中停留足够长的时间从而达到同温层。然而,现在



南极洲臭氧空洞范围在2008年年底达到峰值。

图片来源:NASA

的研究证据表明,该分子能够抵达同温层的较下方,其中包括臭氧层,其高度距离极地8千米。

为了测量二氯甲烷对高海拔臭氧层目前以及未来的威胁,Hossaini 及同事采用了计算

机模拟。2016年,他们的分析表明,南极洲约3%的夏季臭氧流失源自于二氯甲烷。该数量看似较小,但 Hossaini 表示,2010年,该物质对南极洲夏季臭氧流失的1.5%负有责任。研究人员在近日发表于《自然—通讯》的文章中报

告称,如果二氯甲烷排放继续以过去10年的速率增长,那么臭氧空洞的恢复将会推迟30年。

但如果二氯甲烷的排放维持现有水平,臭氧空洞的恢复将会仅仅延后5年左右。不包括二氯甲烷的模拟表明,到2065年,南极洲高海拔地带将会恢复到1980年前的水平,含氯氟烃气体和其他破坏臭氧层的化学物质的浓度则被认为是一个问题。

德国卡尔斯鲁厄理工学院(KIT)大气物理学家 Björn-Martin Sinnhuber 表示,该团队的分析“非常重要”。他指出:“很明确二氯甲烷的浓度已经增加了许多。”但他表示,其中的一个关键问题是长期的排放情况将会如何:“近年来它们变化非常大,很难说它们是如何演变的。”

尽管二氯甲烷排放的迅速增长未来有一天会趋于平稳,但同样,也有可能这些多目的性的化学物质的排放可能会进一步加速。Hossaini 及其团队还评估了如果二氯甲烷排放量以过去20年看到的速率的两倍升高,将会发生什么。其答案会是什么呢?肯定不会好,分析表明,南极洲臭氧层将会在200年以后才能恢复到1980年前的水平。

与此同等重要的是该团队的研究结果可能帮助其他研究人员识别二氯甲烷排量近来的上升的来源。Hossaini 表示,这样的信息目前仍然极为缺乏。(晋楠编译)

科学线人

全球科技政策新闻与解析

特斯拉拟建世界最大“电池”



特斯拉公司首席执行官 Elon Musk

图片来源:Ben Macmahon

近日,美国电动汽车公司特斯拉宣布将在澳大利亚南澳大利亚州打造世界最大的锂离子电池系统以储存电力。这个100兆瓦设备将用于储藏南澳大利亚州风力发电所生产的电能,用于解决电力紧张时的储备问题,这一规模是目前最大蓄电站规模的至少3倍。

“成本效益好的电能储存是阻碍我们从风能和太阳能中获取能量的唯一问题。”澳大利亚格里菲斯大学能源政策专家 Ian Lowe 说。他指出,特斯拉的这个储能系统将“证实大规模储能的可行性”。

澳大利亚能源市场运营机构公布的调查报告指出,南澳大利亚州太阳能和风能生产的电力占比高,电力输出不稳定,需要存储介质才能向电网输入稳定的电流。特斯拉建设的锂电池蓄电站正是用电池作为存储介质,符合南澳大利亚的需求。

悉尼科技大学可再生能源工程师 Geoffrey Jame 也指出,该储能设备将驳斥那些认为可再生能源的可靠性无法与传统化石燃料和核电站的可靠性相匹配的怀疑论者。

在发布会上,南澳大利亚州长杰伊·韦瑟里尔说,特斯拉此举将彻底改变可再生能源存储现状,将能够稳定南澳大利亚电网,控制电价上涨。而且,特斯拉的储能设施将有助于该州向2025年一半能源是可再生能源的方向稳步前进。

据悉,2016年9月,多年不遇的强风和暴雨袭击南澳大利亚,破坏了电力基础设施并导致全州电力供应中断。今年2月,阿德莱德市连续高温导致用电紧张,澳大利亚能源市场运营机构不得不对部分地区断电,使4万民众受到影响。

特斯拉公司首席执行官 Elon Musk 承诺,新蓄电站将在今年12月澳大利亚的夏季开始前投入运行。建成后,这个蓄电站将能够在停电时向3万户家庭供电。(张章)

法科学家因做伪证被判监禁



Michel Aubier

图片来源:CHRISTOPHE ARCHAMBAULT

近日,一场被称为“历史审判”的诉讼在法国一家法院落下帷幕。该法院判决肺病学家 Michel Aubier 因对该国参议院委员会做伪证而被处6个月监禁和5万欧元罚金。

两年前,在调查空气污染的健康风险时,Aubier 曾宣誓自己与业界没有利益冲突。但之后报纸揭露他长期收取石油集团的报酬,参议院因此指控他做伪证。

现年69岁的Aubier 曾任巴黎比沙—克劳德·伯纳德医院肺病部主任,并且还是巴黎第七大学肺学教授。

2015年4月,Aubier 代表巴黎公立医院的一个名为 AP-HP 的研究组,出席了参议院举行的一场旨在了解空气污染的健康影响的听证会。会上,他宣誓称与相关方“没有经济利益纠葛”。

但在2016年3月,法国报纸《解放报》和《鸭鸣报》揭露 Aubier 自1997年开始接受道达尔石油集团报酬,其中包括2012年到2015年间每年超过10万欧元的报酬,以及共计约15万欧元的股份。该集团在接受《解放报》采访时表示,作为回报,Aubier 负责监督公司高管的健康状况并为他们提供公共和职业健康方面的建议。

但 AP-HP 并未察觉这些关联。在许多新闻发布会上,Aubier 通常会发出许多与主流科学观点不同的声音,例如称污染对公众健康影响相当有限。在上述听证会上,Aubier 提出,肺癌与内燃机废气等空气污染的关系“极为微弱”。但法国肺病学会前会长 Bruno Housset 指出,里昂癌症国际研究所早在2012年就将内燃机废气列为致癌物,在2013年将一般空气污染物列入黑名单。“这在科学界是有广泛共识的事情。”Housset 说,“但游说集团常‘兜售’怀疑意见。”

Aubier 并没有出席近日的庭审,并拒绝了《科学》的采访。他的律师 Fran ois Saint-Pierre 则表示,这一审判结果让 Aubier “大受影响”,并认为受到了污蔑。“他在道达尔的医疗顾问工作与他在议会委员会听证会上的行为没有利益冲突,他并没有参与到工业生产中。”Saint-Pierre 说。(张章)