

AI 首次成功从零生成原始蛋白质

本报讯 科学家创建了一个能够从头开始生成人造酶的人工智能(AI)系统。在实验室测试中,尽管人工生成的氨基酸序列与任何已知的天然蛋白质存在显著差异,但其中一些酶与自然世界中发现的酶一样有效。相关研究成果近日发表于《自然-生物技术》。

该实验表明,虽然自然语言处理是为读写语言文本开发的,但至少可以学习一些生物学的基本原理。Salesforce Research 公司开发了名为 ProGen 的 AI 程序,使用下一代标记预测将氨基酸序列组装成人造蛋白质。

科学家表示,这项新技术可能比获得诺贝尔奖的蛋白质设计技术——定向进化更为强大,它将加速新蛋白质的开发,为已有 50 年历史的蛋白质工程领域注入活力。这些新蛋白质几乎可以用于从疾病治疗到降解塑料的任何领域。

“人工设计的性能比受进化过程启发的设计表现得更好。”该研究作者之一、美国加州大学旧金山分校药学院生物工程和治理科学教授 James Fraser 表示,语言模型正在学习进化的各个

方面,但它不同于正常的进化过程。“我们现在能够针对特定效果调整这些属性的生成,例如非常耐热或喜欢酸性环境、不会与其他蛋白质相互作用的酶。”

为了创建这个模型,科学家只需将 2.8 亿种不同蛋白质的氨基酸序列输入机器学习模型,并让它在几周内消化信息。然后,他们使用来自 5 个溶菌酶家族的 56000 个序列,以及有关这些蛋白质的上下文信息对模型进行微调。

该模型迅速生成了 100 万个序列。研究团队根据它们与天然蛋白质序列的相似程度,以及 AI 蛋白质的潜在氨基酸“语法”和“语义”的自然程度,选择了 100 个序列进行测试。

在第一批由 Tierra 生物科学公司体外筛选的 100 种蛋白质中,该团队制作了 5 种人工蛋白质用于细胞测试,并将其活性与鸡蛋清中发现的一种酶(鸡蛋清溶菌酶,HEWL)进行了比较。在人类的眼泪、唾液和牛奶中也发现了类似的溶菌酶,它们可以抵御细菌和真菌。

其中两种人工酶能够分解细菌的

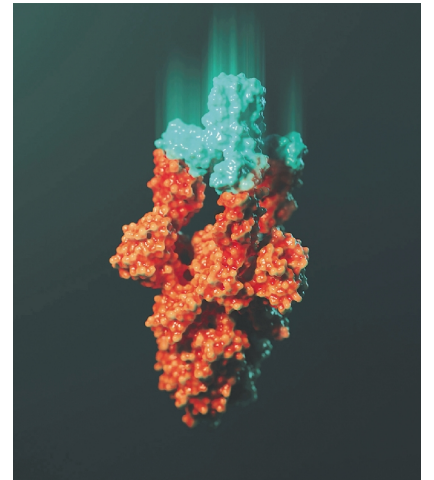
细胞壁,其活性与 HEWL 相当。但它们的序列只有约 18%相同,这两个序列与任何已知蛋白质的相似性分别为 90% 和 70%。

天然蛋白质中的一个突变就能使其停止工作。但在另一轮筛选中,研究小组发现,即使只有 31.4%的序列与任何已知的天然蛋白质相似,AI 生成的酶仍显示出活性。

AI 甚至能够通过研究原始序列数据了解酶如何形成。X 射线晶体学测量显示,人造蛋白质的原子结构看起来和它们应该有的样子一样,尽管这些序列是前所未有的。

2020 年,Salesforce Research 基于研究人员最初开发用于生成英文文本的一种自然语言编程开发了 ProGen。他们从之前的工作了解到,人工智能系统可以自学语法和单词的含义,以及其他使写作井井有条的基本规则。

“当你用大量数据训练基于序列的模型时,它们在学习结构和规则方面确实非常强大,能够了解哪些词可以同时出现,以及组合性。”该研究通讯作者之一、Salesforce Research 人工智能



图片来源:IAN C. HAYDON

研究主管 Nikhil Naik 说。

对于蛋白质,设计选择几乎是无限的。溶菌酶和蛋白质一样小,最多约有 300 个氨基酸。但是有 20 种可能的氨基酸,这样就有 20300 种可能的组合。鉴于无限的可能性,该模型能够如此轻松地产生酶是非常了不起的。

该研究第一作者、Profluent Bio 创始人 Ali Madani 说:“从头开始生成功能性蛋白质的能力表明,我们正在进入蛋白质设计的新时代。对于蛋白质工程师来说,这是可用的多功能新工具,我们期待看到它的治疗应用。” (辛雨)

相关论文信息:<https://doi.org/10.1038/s41587-022-01618-2>

城市树木有助降低高温死亡率

本报讯 一项针对欧洲 93 个城市进行的模拟研究显示,如果将城市树木覆盖率提高 1 倍,夏季因高温死亡的人数会减少近 40%。相关研究近日发表于《柳叶刀》。

欧洲城市的平均树冠覆盖率(从高处俯瞰时覆盖的区域)略低于 15%。英国伦敦和西班牙巴塞罗那的目标分别是到 2030 年和 2037 年将这一覆盖率提高 1 倍,达到 30%。

为搞清实现这一目标的效果,巴塞罗那全球卫生研究所的 Tamara Lungman 等研究人员把 2015 年 6 月至 8 月期间 93 个城市的死亡率与每日温度数据相结合,估算出这 3 个月因高温死亡的人数。

研究人员表示,他们之所以选择研究 2015 年的数据,是因为那一年全欧洲的统计数据最全,且能反映当前欧洲的气候。

随后,研究人员模拟了城市中树木覆盖率增加对温度和死亡率的影响。

该研究共同作者 Mark Nieuwenhuijsen 说:

“人们已经知道树木可以降温。我们准备研究树木究竟能提供多少降温作用,以及能避免多少人死亡。”

2015 年 6 月至 8 月,93 个城市中有 6700 人因极端高温死亡。然而研究表明,如果将森林覆盖率提高到 30%,其中 2644 人(近 40%)可以避免死亡。

树木有助于消除“城市热岛效应”,即城市温度比附近农村地区高。这是因为城市地表,如沥青和混凝土,会吸收并保留热量。在一些地区,城市和附近农村地区的温差可能超过 4°C。

2022 年,英国部分地区温度首次达到 40°C。Nieuwenhuijsen 表示,随着气候变化加快,城市必须做好准备,以应对越来越极端的热浪。

Nieuwenhuijsen 说:“城市中心太热了,我们可以采用基于自然的解决方案,如植树,以减少热岛效应的影响和相关死亡率。” (王见卓)

相关论文信息:[https://doi.org/10.1016/S0140-6736\(22\)02585-5](https://doi.org/10.1016/S0140-6736(22)02585-5)

本报讯 美国麻省理工学院 Elly Nedivi 研究组将丘脑神经支配映射到单个 L2/3 锥体神经元并通过建模“读出”视觉输入。相关论文近日发表于《自然-神经科学》。

丘脑是感觉信息从外周到哺乳动物大脑皮层的主要通道。作为进入新皮层的主要前馈投射系统,丘脑的核心作用与丘脑皮层突触的疏密相关。

研究人员采用新方法,结合遗传工具和可伸缩的组织扩展显微镜进行全细胞突触映射,揭示了主视觉皮层的 L2/3 锥体细胞(PC)上的丘脑和皮质兴奋性突触的数量、密度和大小。研究人员发现丘脑的输入不仅稀疏,而且

在单个树突和神经元之间的数量和密度上也存在显著的异质性。最令人惊讶的是,尽管 L2/3 PC 上的丘脑突触稀疏,但它们比皮层上的突触要小。研究人员将这些发现纳入精细尺度、解剖学上可靠的 L2/3 PC 生物物理模型,揭示了整合在小的异质神经元集合中的具有稀疏和微弱丘脑皮层突触的单个神经元能可靠地“读出”视觉驱动的丘脑输入。 (柯讯)

相关论文信息:

<https://doi.org/10.1038/s41593-022-01253-9>

科学家用建模「读出」视觉输入