人工智能破解蛋白质复合物密码

作者：writer 来源：爱科学

本文原地址：https://www.iikx.com/news/progress/16591.html

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

人工智能破解蛋白质复合物密码。在蛋白质结构预测上，人工智能革命仍在继续。一年前，软件程序首次成功地模拟了单个蛋白质的3D形状，其精确度与几十年前的实验技术测出的一样准确。今年夏天，研究人员利用人工智能程序编程了一个近乎完整的人类蛋白质结构目录。

现在，美国研究人员更进一步，使用人工智能技术确定了不同蛋白质之间可能的相互作用，以及由这种相互作用产生的复合体是什么样子的。相关研究结果发表于《科学》。

斯坦福大学系统生物学家Michael Snyder表示，这项新发现会促进细胞生物学领域产生一系列新观点，并为发现下一代治疗药物指明新方向。

几十年来，精确重现人类蛋白质的形态和结构，一直是研究人员的目标。过去，要追求这个目标是很困难的，需要昂贵且缓慢的实验，如X射线晶体学和核磁共振光谱学实验。而且，即使实验取得进展，也只能得到单个蛋白质某些结构。

计算机专家一直致力于加快实现该目标的速度。借助人工智能和深度学习算法，使用实验解决的蛋白质结构数据库来训练软件程序如何根据蛋白质的氨基酸序列预测蛋白质的结构。

该研究同时使用了华盛顿大学的RoseTTAFold和DeepMind的AlphaFold工具，筛选了830万对酵母蛋白的配对多序列比对，识别出了1505种可能的相互作用，并为106个以前未识别的部分和806个尚未结构表征的部分构建了结构模型。

为了找到可以形成复合物的蛋白质，研究团队首先将所有6000种酵母蛋白质的氨基酸序列，与2026种其他真菌、4325种其他真核生物的氨基酸序列进行比较。在比较过程中，研究人员追踪这些蛋白质的进化过程，并识别出不同蛋白质中同时发生变化的序列。研究人员推断，这些蛋白质可能会形成复合物，并会逐步改变以保持它们之间的相互作用。

随后，团队使用自研的RoseTTAFold的人工智能程序以，以及此前DeepMind已开源的 AlphaFold工具，尝试解决每组候选对象的3D结构。在总共830万个候选酵母蛋白质对中，两个人工智能程序合力识别出了1506种可能发生相互作用的蛋白质，并成功绘制出其中712个的3D结构。

研究通讯作者之一、德克萨斯大学西南医学中心生物医学信息学专家Qian Cong表示，这些相互作用涵盖了真核细胞的所有活动。该研究的亮点在于发现了蛋白质复合物的结构，这种复合物在细胞活动中发挥着不可替代的重要作用，如让细胞修复DNA损伤，将RNA翻译成核糖体中的蛋白质，在细胞繁殖过程中将染色体拉开，并将分子运送通过细胞膜等。

AlphaFold首席开发人员之一、DeepMind团队成员John Jumper也对这一发现的意义给予了高度评价：这是再现蛋白质3D结构的一个很有前途的研究实例。明确了蛋白质之间如何相互作用，生物学家就可以进一步发现复合物细胞内执行多项任务的机制。

这些模型为实验人员提供了可测试的假设，而且破坏这些相互作用，可能会获得干预各种疾病的新方法。Qian Cong表示，这一发现为未来新靶向药物的研发提供了更多可能。

上个月，Jumper团队在bioRxiv上发布了一篇预印本论文，介绍了AlphaFold人工智能工具的新版本，名为AlphaFold-Multimer。该工具绘制出了4433种蛋白质复合物的结构，准确率达到了69%。（来源：中国科学报辛雨）

相关论文信息：https://doi.org/10.1126/science.abm4805

版权声明：凡本网注明来源：中国科学报、科学网、科学新闻杂志的所有作品，网站转载，请在正文上方注明来源和作者，且不得对内容作实质性改动；微信公众号、头条号等新媒体平台，转载请联系授权。邮箱：shouquan@stimes.cn。
作者：Qian Cong 来源：《科学》

更多 科学进展 请访问 https://www.iikx.com/news/progress/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发