一类神经网络具有了类人的泛化语言能力

作者：writer 来源：科学网

本文原地址：https://www.iikx.com/news/progress/24725.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！

一类神经网络具有了类人的泛化语言能力。科学家创造了一种神经网络，它具有类似人类的能力，可以对语言进行概括。它在将新学到的单词折叠到现有词汇表中并在新的语境中使用方面的表现与人类不相上下，这是人类认知的一个关键方面，被称为系统泛化。

研究人员将同样的任务交给了聊天机器人ChatGPT的AI模型，发现它在这样的测试中的表现比新的神经网络或人要差得多，尽管聊天机器人具有以类似人类的方式进行对话的神奇能力。

这项工作可能会使机器与人的互动比当今最好的人工智能（AI）系统做的更自然。相关成果10月25日发表于《自然》。

美国约翰斯·霍普金斯大学专门研究语言的认知科学家Paul Smolensky说，神经网络的类人表现表明，使网络具有系统性的训练能力方面已经取得了突破。

系统泛化能力表现为人们在新环境中毫不费力地使用新获得的单词。例如，一旦人们掌握了照片炸弹一词的含义，他们就可以在各种情况下使用它；同样，理解猫追狗的人不需要太多思考也会理解狗追猫。

但是论文作者之一、美国纽约大学认知计算科学家BrendenLake说，神经网络并非天生具备这种能力。神经网络是一种模拟人类认知的方法，一直主导着AI研究。与人类不同的是，神经网络很难使用一个新词，直到它们在使用该词的许多样本文本上进行了训练。近40年来，AI研究人员一直在争论，如果神经网络不能证明具备这种能力，它能否成为一种合理的人类认知模型。

为结束这场争论，研究人员首先测试了25个人，看他们在不同情况下使用新学单词的能力。他们使用一种由两类无意义单词组成的伪语言来测试，以确保参与者是首次学习这些单词。daxwif和lug等原始单词代表了跳过和跳跃等基本而具体的动作。更抽象的功能单词，如blicketkiki和fep，则指定了使用和组合的规则，从而产生了跳三次或向后跳之类的组合。

参与者被训练将每个原始单词与特定颜色的圆圈联系起来，因此红色圆圈代表dax，蓝色圆圈代表lug。然后，研究人员向参与者展示了原始单词和功能单词的组合，以及当后者应用于前者时产生的圆圈模式。最后，研究人员通过向参与者提供复杂组合来测试他们应用这些抽象规则的能力。参与者必须选择正确的颜色和圆圈数量，并将其按适当的顺序放置。

正如预测的那样，人们在这项任务上表现出色，平均有80%的时间选择了正确的彩色圆圈组合。当他们犯错误时，研究人员注意到这些错误遵循了一种反映已知人类偏见的模式。

接下来，研究人员训练了一个神经网络，通过编程让它从错误中学习，从而完成与提供给参与者的任务类似的任务。这种方法允许AI在完成每项任务时进行学习，而不是使用静态数据集——这是训练神经网络的标准方法。为了使神经网络像人类一样，研究人员训练它重现在人类测试结果中观察到的错误模式。当神经网络进行新测试时，它的答案几乎与人类参与者的答案完全一致，在某些情况下甚至超过了人类的表现。

相比之下，GPT-4在完成同样的任务时遇到了困难，平均失败率在42%到86%之间——取决于研究人员如何呈现任务。这不是魔法，而是练习。Lake说，就像孩子在学习母语时也会练习一样，这些模型通过一系列学习任务来提高它们的技能。

美国圣达菲研究所的计算机和认知科学家MelanieMitchell表示，这项研究是一个有趣的原理证明，但这种训练方法能否扩展到更大的数据集甚至图像上，还有待观察。

Lake希望通过研究人们如何从小就培养出系统泛化的能力来解决这个问题，并将这些发现结合起来，建立一个更强大的神经网络。（来源：中国科学报 文乐乐）

相关论文信息：https://doi.org/10.1038/s41586-023-06668-3

作者：Paul Smolensky 来源：《自然》

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