专家解读：何谓大语言模型？

作者：杨东 来源：爱科学

本文原地址：https://www.iikx.com/news/topnews/24907.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！

大语言模型(英文Large Language Model，简称LLM)，也称大型语言模型，是指能够处理和生成自然语言的计算机模型；它代表着人工智能领域的重大进步，并有望通过习得的知识改变该领域。在过去几年中，LLM的规模每年增加10倍，而且随着这些模型的复杂程度和规模的增加，其性能也在不断发展。



LLM的特点是规模庞大，包含数十亿的参数，帮助它们学习语言数据中的复杂模式;这些复杂模式通常基于深度学习架构(如转化器)，这有助于它们在各种自然语言处理任务上取得令人印象深刻的表现。拿GPT来说，GPT 其实出现了好几代，其中GPT-3有45个TB的训练数据，那么整个维基百科里面的数据只相当于它训练数据的0.6%。

这些模型也就是语言材料，这个语料的量可以说是集中到我们人类所有语言文明的精华在里面，这是一个非常庞大的数据库。LLM经过这样的一个量的学习之后，它产生的数据量就是一些做人工智能技术的也没有想到会有这种变化——当数据量超过某个临界点时，模型实现了显著的性能提升，并出现了小模型中不存在的能力，比如上下文学习(in-context learning)。

LLM可以通过学习语言数据的统计规律和语义信息来预测下一个单词或句子，随着输入数据集与参数空间的不断扩大，LLM的能力也会相应提高。它用于多种应用领域，如机器人学、机器学习、机器翻译、语音识别、图像处理等，所以被称为多模态大型语言模型(MLLM)。

MLLM是在LLM的基础上，融合了其他类型的媒体数据(如图像、视频、声音)。它能够同时处理不同媒体类型的信息，并将其整合到一个统一的语义空间中。这使得它在理解多媒体数据的任务中具有很强的优势，如图像描述、视频理解、多模态问答等。

相对于传统的单模态处理方法，MLLM可以更好地理解和描述复杂的现实世界中的信息。它通过将不同媒体类型的信息进行编码和融合，能够更准确地捕捉多媒体数据中的语义和上下文信息。在自然语言处理方面，LLM已经成功地在各种自然语言任务中充当通用接口。

前不久，由美国谷歌公司和德国柏林工业大学的人工智能研究团队推出了史上最大的视觉语言模型(PaLM-E)，该模型的参数量高达5620亿(GPT-3的参数量为1750亿)。作为至今最强大脑的PaLM-E不仅可以理解图像，还能理解、生成语言，执行各种复杂的机器人指令而无需重新训练；同时，它还展示出了强大的涌现能力。

最近，美国微软公司的研究团队发表了一篇名为《机器人ChatGPT：设计原则和模型能力》论文，公布了他们正在把ChatGPT应用于机器人上的研究成果。这项研究的目标是观察ChatGPT是否可超越文本思考，并对物理世界进行推理来帮助完成机器人任务;在监控机器人性能的同时向LLM提供高级反馈。

目前中国在LLM方面的创业公司相对IT大厂来说主要有两个优势，一是技术，另一是数据。由于他们已经推出了相关产品，“ 数据飞轮 ”已经转起来了，这些数据的质量相比互联网数据质量要高很多，能够对产品迭代起到很大支撑作用。

从2022年底开始，LLM在全球掀起热潮，人工智能技术的发展获得里程碑式的突破，并产生了广泛的影响。正如中国著名学者周海中先生在20世纪90年代初所言：“ 人工智能技术将广泛应用到各学科领域，会产生意想不到的效果。 ”

可以预料，随着人工智能技术的不断发展和应用场景的不断扩大，LLM未来将在更多领域得到应用和推广，它将为人类带来更加美好的未来。

文/杨东(作者单位：美国布朗大学工程学院)

更多 科研头条 请访问 https://www.iikx.com/news/topnews/

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