

李飞飞、吴恩达对谈：这一次，AI 冬天不会到来

01

AI 会继续深化、泛化

Q：AI 已经经历了好几轮热潮和低谷，今年 AI 会怎么样？会迎来低谷吗？

吴恩达：我觉得不会，因为人工智能的商业基础比以往任何时候都更为强大，甚至在生成式人工智能浪潮之前就已经如此。**它们在去年真的起飞了，人工智能可能已经推动了数千亿美元，也许是数万亿美元。**至少对于像谷歌这样的公司来说，它给你展示了更相关的广告，从而带来了大量收入。因此，商业基础就在那里。

事实上，人工智能最为人所不理解的一个方面是它是一种通用技术，这意味着它不仅对某一方面有用，这有点类似于电力。如果我问你电力有什么用，几乎很难回答，因为它对许多不同的事物都有用。

人工智能也是如此。所以今天即使人工智能还没有取得技术上的核心进展，但你知道它会取得技术进展。即使没有，全球各地已经有太多场景等待被发现和建设，（人工智能的）商业基础会继续增长。

李飞飞：我大致同意恩达的观点。我们所看到的是**人工智能的一个转折点，这个转折点是随着通过大型语言模型、ChatGPT 以及随后的模型的出现而实现的。**

我看到，人工智能正在成为一种深化的横向技术。当它成为一种深化的横向技术时，它就会成为**下一次数字革命或工业革命真正的变革性驱动力。**这项技术已经来临，正在深入到所有垂直业务和客户、消费者体验中，并正在改变我们社会、经济、政治景观的基本结构。事实就是这样，而且这种变化正在变得越来越明显。

Q：2024 年，人工智能领域可能发生的重大突破是什么？

吴恩达：首先，我看到了**从大型语言模型向大型视觉模型的转变。**很多进展不仅仅是在生成图像方面，还包括对图像进行分析，以便计算机能够更好地看清图像的含义，对于自动驾驶等领域具有重要影响。

我对自主代理（Auto Agents）的崛起感到兴奋，这是指你可以给一个人工智能系统一个指令，比如，「**亲爱的人工智能系统，请帮我进行市场调研，进行对该公司的竞争分析。**」**它会规划一系列动作，如网络搜索、下载网页并总结信息。**它会离开一会，进行半小时、一个小时或一天的工作，然后回来提供答案。这些代理可以规划和执行一系列动作。自主代理现在只是刚刚有点起色，但我觉得在研究和商业化方面有很大潜力。我期待在未来几个月内会有突破性进展。

最后一点是边缘人工智能（edge AI）。我经常在我的笔记本上运行大型语言模型，我一直在使用 GPT-4，但是有很多 bug。但很多人可能不知道的是，实际上在你的笔记本上运行大型语言模型已经变得相当可行，虽然不像 GPT-4 那么大，但足够大、足够有用。

Q：在你的笔记本电脑、个人电脑或工业 PC 上运行边缘人工智能，这方面的能力实际上比大多数人认为的要好得多。我认为这将推动很多设备的销售，也将推动很多计算机科学的发展。

李飞飞：我将以尊重的态度稍微持不同意见，或者说我们可以进行一点讨论。

Q：你用了自主代理（Autonomous Agent）这个词，我实际上想将它改成辅助代理（assistive agent）。

我们在今天的大语言模型中看到的一件事是，**长尾分布**仍然很困难，无论我们谈到的是「幻觉」还是其他事情。在很多工作领域中，为了提供高质量的服务和产品，长尾问题是很重要的。所以我实际上看到的是一种人机协同的、辅助代理的情况，其中一部分工作是自主完成的，一部分工作是协作完成的。这是更有可能的情况，而不是完全自主。

吴恩达：我们终于有点不同意见了，不过实际上我是同意你的。让我分享一下我的经验。

我认为「自主代理」这个术语可能有问题，但在商业世界中我看到的是，**很多人宁愿让人工智能与人类共事，而不是取代人类，因为这会引发有关失业的讨论。**在我的观察中，**是否使用人工智能进行自动化或削减人手，通常是一个商业经济决策，而不是伦理决策。**也许它应该是一个伦理决策，但坦率地说，在我与企业合作构建聊天机器人时，我看到大多数企业进行了非常理性的经济计算，因为竞争对手也在做同样的事情。

李飞飞：我要补充一点，然后对此进行评论。

我认为我们必须小心**代替工作**（job）和**代替任务**（task）的区别。每一个人类工作，实际上是涉及到多个任务的集合。比如我研究医疗保健，一个护士 8 小时的工作包含数百个任务。我确实看到人工智能代理在许多任务上提供帮助、具有辅助和增强功能，但在讨论工作时我们应该非常小心。

我认为经济上的商业决策与伦理社会决策并不是互相排斥的。这是一个更深入的对话，至于你提到的这些代理是否具有理解能力（understanding），仅就商业而言，我认为这是一个非常微妙的术语。什么是理解？有在数据中存在的模式的理解，有关于你正在做出的决策的理解，还有关于人类任务意图的理解。所以我认为实际上我不会那么过于笼统地使用「理解」这个词来描述今天的人工智能代理。

吴恩达：我想回到任务的话题，因为我认为这很重要。

我的团队为许多企业工作，有时我会接到一位首席执行官的电话。他们说：「嘿，恩达，我在阅读关于人工智能代理的文章，我该怎么办？」

我们的朋友艾弗里·布朗（Avery Browne）经常给出的建议，就是观察你的团队，**找出你所有员工实际在做哪些任务，并分析不是在工作层面而是在任务层面上**，这些任务对于人工智能增强或自动化来说有多容易，并且其业务回报率是多少。每次我在企业中进行这项工作时，我们总是想出非常多的想法。因此，人工智能增强或自动化有很多机会。

第二件事是，往往具有最高回报率的任务并不是人们最初认为的那样。例如，当你想到放射科医生时，人们往往认为，哦，放射科医生要看 X 光片。这是你脑海中有这份工作的认知，但实际上我们将这份工作分解为许多不同的任务，比如收集患者病史之类的事情，结果可能更容易，而且回报率更高。

因此，我发现系统地进行这种练习通常有助于企业识别有价值的机会。

李飞飞和吴恩达在 CES 展会论坛上畅谈接下来 AI 和大模型的发展和应用 | CES

■AI 工具：ChatGPT、Dall-E、Claude、Midjourney、Stable Diffusion 等使用入口



■AI 行业研报、书籍、论文（持续更新）

①AI 学社资源目录 ↓



②AI 学社入口 ↓



■AI 工具代理副业

①AI 工具代理副业介绍 ↓



②客服微信 ↓





02

在不确定性中 寻找确定性

Q：在你与财富 500 强公司合作的工作中，有没有一些应用领域具有清晰、可证明、可实现的投资回报率（ROI）？

吴恩达：如果我们从最广泛的角度来看财富 500 强公司，**我认为常见的领域是客户运营或客户支持，有很多公司试图增强或自动化客户支持。**销售运营、软件工程也受到了很大的影响。

但对我来说更令人兴奋的是，我正在与一个非常大的农业公司讨论，我们认为也许可以使用人工智能来执行真正为他们节省很多时间的奇怪任务。所以那些具体到你的业务和你的行业的细分市场，我认为通常更有趣，可以创造出特定于行业的防御性的强大的增长机制。人们在某个时候可能都会购买一些用于销售运营等方面的通用工具，但是对于你的业务而言，那些应该在内部构建的具体事物，实际上是非常令人兴奋。

李飞飞：有各种各样的客户支持或运营解决方案，要补充的一点是，当前技术中存在哪些共同的机会？我认为这在今天仍然是真实的，就是**在你拥有最多数据的地方，在数据能够显示出可重复的模式或好的模式的地方，就是你可以开始的地方。**无论是人类语言模式、结构化数据模式还是图像数据模式，数据存在的地方以及数据模式被证明在你的业务中是有价值且可操作的地方，那就是值得关注的地方。

Q：这些财富 500 强公司可能面临的障碍是什么？许多 CEO 担心 AI 产生的不确定性。

李飞飞：这取决于你的产品、服务，取决于结果的风险。在医疗、金融预测等许多行业中，准确性对长尾影响如此之大，以至于你不能承担人身伤害的风险，或者金融错误的风险。所以，你需要了解你的行业，你需要了解你的解决方案和服务。

这就是为什么我们**应该远离炒作，理解这项技术能做什么，并避免在那种 AI 尚未准备好的方向投资。**

极客公园

，赞251

Q：顺便说一句，生成式 AI 是一个「负担过重」的词

。今天每一个 AI 都被称为生成式 AI，但是在一开始的时候，我们对生成式 AI 有非常明确的[数学定义](#)。

确切地说，当谈到生成式 AI 时，我只是假设它是一种大规模的数据驱动，有一个预训练阶段的方法。有些人可能会将 Transformer 模型和预测模型放入其中，但无论如何，我认为如果存在真正的准确性问题，我们应该检查几个方面：**这是模型的限制吗？这是数据质量的问题吗？AI 在正确的循环中吗？**有更微妙的业务问题导致了不准确。真的要弄清楚所有这些，并尽力解决它们。有时，例如在某些级别的医疗诊断和治疗中，你必须认识到存在一些限制，如果风险太高，我们不能过于激进。

Q：让我们讨论一下基础模型。如果 2023 年是 OpenAI 的一年，那么在基础模型领域，谁将是我们 2024 年最多谈论的人？苹果会推出 Ajax LLM 吗？我们会更多地谈论 Gemini 而不是 GPT 吗？

李飞飞：我之前说过，我看到这项技术正在深化，并且也在扩展到所有领域。因此，很难单独挑出一个公司。我相信将会有令人振奋的发布，从芯片方面的下一代 H100（英伟达 AI 专用芯片组）的发布，一直到消费端。

因此，我无法押注一个单一的话题，但我希望看到，我也相信，**2024 年将会是由 AI 应用和 AI 技术拓宽定义的一年。**

吴恩达：每当有一波技术创新时，媒体喜欢谈论技术层面，这就是为什么媒体关注 OpenAI、Google、AWS、Microsoft、Meta、Nvidia 和 AMD 等的原因。这没有错，但事实证明，**为了使这个技术基础设施层取得成功，我们需要另一个领域取得更大的成功，那就是建立在这些云服务提供商之上的应用层。**

因为坦白说，我们需要建立在这些工具之上的应用来创造更多的收入，以便它们能够支付这些工具的构建者的成本。我们最好集体创造一些生成应用，以填补 GPU 资本投资带来的数百亿美元的空缺。因此，我认为很多实际工作不仅仅是基础模型。它将涉及到医疗、金融服务、教育、电子商务等各种行业。

03

AI 版权问题

Q：今天的一个主要话题是关于生成式人工智能和知识产权的诉讼。你如何看待这些诉讼的发展？《纽约时报》是否应该因 OpenAI 在训练中使用其内容而获得补偿？

吴恩达：我查看了《纽约时报》、OpenAI 和微软的诉讼，我不是律师，也不提供法律建议或任何有关此事的任何建议，但我对 OpenAI 的同情要远远超过对微软的同情，我对 OpenAI 和微软的同情也要远远超过对《纽约时报》的同情。

坦率地说，当我看到《纽约时报》的诉讼时，我觉得这是一个非常混乱的论点。**我希望《纽约时报》的律师能够遵循与记者相同的清晰和新闻报道标准，但我认为他们没有这样做。**

简而言之，有两件事。

首先是关于提示（prompt）的问题，即输入给 OpenAI 来「复读」（regurgitate）《纽约时报》文章的指令。我认为这是一个非常奇怪的指令，几乎没有任何正常用户会使用。我不认为大规模复读受版权保护的文章是应该的。但是 OpenAI 存在这个 bug，而《纽约时报》只是指出了这个 bug。我们都知道，有时在软件中会出现 bug。

还有另一件奇怪的事情，在一些例子中，《纽约时报》展示了你可以编写一个提示，将其传递给 ChatGPT，要求其下载《纽约时报》的文章，然后告诉它将其打印出来。

我觉得仅仅因为它能够这样做，这并不等同于 OpenAI 训练了大量的文本数据，包括《纽约时报》的文章。我认为这个诉讼试图在「OpenAI 训练了包括《纽约时报》文章在内的大量文本数据的事实」与「OpenAI 正在进行大规模复读《纽约时报》文本的观点」之间建立联系，我认为这并没有完全反映事实的全部。

李飞飞：我确实想要补充一点，并从更宽泛的角度谈谈生成式人工智能与创作者经济之间的紧张关系。

虽然我在《纽约时报》的争端方面并不像恩达那样了解，但即使在我的书中，我也提到了这项技术的混乱性。我们这些在技术领域受过训练的人喜欢看到确定性。但事实是，**当技术真正应用时，特别是像这样深刻的技术，它就会与人类世界、人类社会变得复杂起来。**

《纽约时报》与 OpenAI 和微软的诉讼实际上显示了我们在生成式人工智能技术与创作者经济之间所看到的紧张关系。互联网已经迅速推动了创作者经济，这影响不仅仅是像纽约时报这样的大公司，还有像单个艺术家、摄影师、音乐作曲家这样的小公司，整个生态系统也受到了挑战、扰乱，以及被今天的生成式人工智能技术所增强。除了纽约时报的诉讼之外，我们还看到了艺术家与 Midjourney 和其他公司的诉讼。

所以我也请媒体以更加细致入微的视角来看待这个问题。无论是公共部门还是私营部门，我们都应该更深入地关注这个问题，而不仅仅是挖掘表面。

吴恩达：在好莱坞罢工期间，许多困难之一就是，如果你是一名创作者，你会想，我的工作会不会消失？我的所有作品还会存在吗？我对此深感同情。

我认为对于失业的担忧可能比实际情况更为严重，因为正如我们之前所说，工作是由任务构成的，即使是艺术家的工作也是由许多任务构成的。是的，人工智能可以自动完成一些任务，但仍然有很多其他任务需要人来完成，而且他们可能会更加高效，实际上还能赚更多的钱。

但有这种担忧，我认为这是一个挑战。我认为人工智能领域需要在这方面做得更好，进行更好的对话，并向人们保证工作丧失的情况不会像他们想象的那么糟糕。

顺便说一句，**版权法是在先前的时代制定的，我认为它需要为生成式 AI 时代进行更新。**