

# 第一章 简介

欢迎来到**面向开发者的提示工程**部分，本部分内容基于**吴恩达老师**的《**Prompt Engineering for Developer**》课程进行编写。《Prompt Engineering for Developer》课程是由**吴恩达老师**与 OpenAI 技术团队成员 **Isa Fulford** 老师合作授课，Isa 老师曾开发过受欢迎的 ChatGPT 检索插件，并且在教授 LLM（Large Language Model，大语言模型）技术在产品中的应用方面做出了很大贡献。她还参与编写了教授人们使用 Prompt 的 OpenAI cookbook。我们希望通过本模块的学习，与大家分享使用提示词开发 LLM 应用的最佳实践和技巧。

网络上有许多关于提示词（Prompt，本教程中将保留该术语）设计的材料，例如《30 prompts everyone has to know》之类的文章，这些文章主要集中在 **ChatGPT 的 Web 界面**上，许多人在使用它执行特定的、通常是一次性的任务。但我们认为，对于开发人员，**大语言模型（LLM）的更强大功能是通过 API 接口调用，从而快速构建软件应用程序**。实际上，我们了解到 DeepLearning.AI 的姊妹公司 AI Fund 的团队一直在与许多初创公司合作，将这些技术应用于诸多应用程序上。很兴奋能看到 LLM API 能够让开发人员非常快速地构建应用程序。

在本模块，我们将与读者分享提升大语言模型应用效果的各种技巧和最佳实践。书中内容涵盖广泛，包括软件开发提示词设计、文本总结、推理、转换、扩展以及构建聊天机器人等语言模型典型应用场景。我们衷心希望该课程能激发读者的想象力，开发出更出色的语言模型应用。

随着 LLM 的发展，其大致可以分为两种类型，后续称为**基础 LLM**和**指令微调（Instruction Tuned）LLM**。**基础 LLM**是基于文本训练数据，训练出预测下一个单词能力的模型。其通常通过在互联网和其他来源的大量数据上训练，来确定紧接着出现的最可能的词。例如，如果你以“从前，有一只独角兽”作为 Prompt，基础 LLM 可能会继续预测“她与独角兽朋友共同生活在一片神奇森林中”。但是，如果你以“法国的首都是什么”为 Prompt，则基础 LLM 可能会根据互联网上的文章，将回答预测为“法国最大的城市是什么？法国的人口是多少？”，因为互联网上的文章很可能是有关法国国家的问答题目列表。

与基础语言模型不同，**指令微调 LLM**通过专门的训练，可以更好地理解并遵循指令。举个例子，当询问“法国的首都是什么？”时，这类模型很可能直接回答“法国的首都是巴黎”。指令微调 LLM 的训练通常基于预训练语言模型，先在大规模文本数据上进行**预训练**，掌握语言的基本规律。在此基础上进行进一步的训练与**微调（finetune）**，输入是指令，输出是对这些指令的正确回复。有时还会采用**RLHF（reinforcement learning from human feedback，人类反馈强化学习）**技术，根据人类对模型输出的反馈进一步增强模型遵循指令的能力。通过这种受控的训练过程。指令微调 LLM 可以生成对指令高度敏感、更安全可靠输出，较少无关和损害性内容。因此。许多实际应用已经转向使用这类大语言模型。

因此，本课程将重点介绍针对指令微调 LLM 的最佳实践，我们也建议您将其用于大多数使用场景。当您使用指令微调 LLM 时，您可以类比为向另一个人提供指令（假设他很聪明但不知道您任务的具体细节）。因此，当 LLM 无法正常工作时，有时是因为指令不够清晰。例如，如果您想问“请为我写一些关于阿兰·图灵(Alan Turing)的东西”，在此基础上清楚表明您希望文本专注于他的科学工作、个人生活、历史角色或其他方面可能会更有帮助。另外您还可以指定回答的语调，来更加满足您的需求，可选项包括**专业记者写作**，或者**向朋友写的随笔**等。

如果你将 LLM 视为一名新毕业的大学生，要求他完成这个任务，你甚至可以提前指定他们应该阅读哪些文本片段来写关于阿兰·图灵的文本，这样能够帮助这位新毕业的大学生更好地完成这项任务。本书的下一章将详细阐释提示词设计的两个关键原则：**清晰明确**和**给予充足思考时间**。