



利用提示工程构建 大语言模型 (LLM) 应用



课程概述

大语言模型 (LLM) 具有强大的能力，全球的企业机构都在尝试将其整合到产品和内部应用中，例如文本生成、大规模文档分析和聊天机器人助手等等。

在各种应用中融入 LLM 的最快方法是通过使用先进的提示工程技术，这也是相对于检索增强生成 (RAG)、参数高效微调 (PEFT) 等更高级 LLM 技术的基础技能。在本课程中，您将使用由开源 Llama-3.1 大语言模型驱动的 NVIDIA NIM™ (一套用于部署 AI 模型的推理微服务，是一个预构建容器工具)，和当前流行的 LangChain 库，学习使用提示工程一系列基本技能，以用于构建基于 LLM 的应用。

学习目标

参加本次培训，您将能够：

- > 理解如何应用迭代式提示工程最佳实践，以创建基于 LLM 的应用程序，用于各种与语言相关的任务
- > 熟练使用 LangChain 来组织和编排 LLM 工作流
- > 编写应用代码，利用 LLM 执行生成任务，比如文档分析、聊天机器人等

培训详情

课程时长	8 小时，课后 6 个月内可以继续访问课件（实验资源用量有限额）
课程模式	讲师实时授课，每位学员可使用完全配置的云端实验环境进行练习
预备知识	> 具备 Python 编程经验，对 LLM 基础知识有扎实的理解，推荐学习《构建基于大语言模型 (LLM) 的应用》
使用的工具、库和框架	NVIDIA NIM, LangChain, Llama 3.1
学员评测方式	基于技能的编码项目，挑战学生编写各种基于 LLM 的应用代码的能力
培训证书	成功完成本课程和测试后，学员将获得 NVIDIA DLI 培训证书，证明在相关领域的的能力，为职业发展提供证明
课程语言	中文
学习此课程的硬件要求	您需要一台笔记本电脑或台式机，且能够运行最新版 Chrome 或 Firefox 浏览器。我们为您提供在云端完全配置的加速工作站的专用访问权限
课程价格	> AI 培训班：每人 3500 元 (提供发票) > 企业专属培训：联系我们，微信添加 NVIDIALearn

为何选择 NVIDIA 深度学习培训中心 (DLI) 的实战培训?

- > 学习 NVIDIA 与技术专家和行业领导者合作开发的课程, 获取全球同步、技术领先和现实可用的专业开发技能和经验。
- > 学习使用行业通用、标准的软件、工具和框架进行端到端的应用开发, 能够在广泛的行业中构建基于深度学习、加速计算、图形与仿真和数据科学的应用。
- > 系统化地学习理论, 并使用云端完全配置的实验环境同步边练, 高效提升实战开发能力。
- > 获得 NVIDIA 全球开发者培训证书, 加持专业培训认证, 助力职业发展。

课程大纲

课程介绍 (15 分钟)	<ul style="list-style-type: none">> 讲师介绍> 登录 DLI 学习平台
引言 (45 分钟)	<ul style="list-style-type: none">> 了解为什么提示工程是与大语言模型 (LLM) 交互的核心> 讨论如何使用提示工程来开发多种基于 LLM 的应用> 了解 NVIDIA LLM NIM, 它被用于部署本课程中使用的 Llama 3.1 LLM
提示词入门 (60 分钟)	<ul style="list-style-type: none">> 熟悉课程环境> 使用 OpenAI API 和 LangChain 创建并查看你的第一个提示响应> 学习如何流式传输 LLM 响应, 以及批量发送 LLM 提示, 比较性能差异> 练习迭代式提示开发的过程> 创建并使用你的第一个提示模板> 完成一个小项目, 在一批输入上执行分析和生成任务
休息 (60 分钟)	
LangChain 表达式语言 (LCEL), 运行时, 链 (60 分钟)	<ul style="list-style-type: none">> 了解 LangChain 运行时, 以及用 LangChain 表达式语言 (LCEL) 将它们组合成链的能力> 编写自定义函数并将其转换为运行时, 以包含进 LangChain 链中> 将多个 LCEL 链组合成一个更大的应用链> 通过组合并行 LCEL 链来发挥并行工作的能力> 完成一个小项目, 借助 LCEL 和并行执行, 对一批输入执行分析和生成任务
带消息的提示词 (45 分钟)	<ul style="list-style-type: none">> 了解两种核心聊天消息类型: 人类消息和 AI 消息, 以及如何在应用代码中明确地使用它们> 通过一种被称为少样本提示的技术, 为聊天模型提供指导性示例> 明确使用系统消息, 这将允许您为聊天模型定义一个总体的角色和职责> 使用思维链提示来增强 LLM 执行复杂推理任务的能力> 管理消息以保留对话历史, 构建聊天机器人> 完成一个小项目, 构建一个简单但灵活的聊天机器人应用, 能够担任各种角色
休息 (15 分钟)	
结构化输出 (60 分钟)	<ul style="list-style-type: none">> 探索一些基本方法, 使用 LLM 批量生成结构化数据以供下游使用> 结合 Pydantic 类和 LangChain 的 `JsonOutputParser` 生成结构化输出> 学习如何从长篇文本中提取数据和进行标记> 完成一个小项目, 使用结构化数据生成技术对非结构化文本文档执行数据提取和文档标记
使用工具、构建智能体 (75 分钟)	<ul style="list-style-type: none">> 创建一个能够推理何时适合使用工具, 并将工具使用结果整合到其响应中的智能体> 完成一个小项目, 创建一个 LLM 智能体, 能够利用外部 API 调用来增强其响应, 提供实时数据

课程大纲

评估测试和总结

(30 分钟)

- > 回顾所学要点并回答问题
- > 完成评估并获得证书
- > 填写培训调查表
- > 获得进一步学习的建议

下一步

学习更多 DLI 课程：

- > 高效定制大语言模型 (LLM)
- > 构建基于大语言模型 (LLM) 的应用
- > 构建基于扩散模型的生成式 AI 应用
- > 模型并行 —— 构建和部署大型神经网络

准备好开始学习了吗？

查询更多 DLI 课程，请访问 nvidia.cn/training

如有疑问，请通过微信联系 DLI 小助手 (微信号 NVIDIALearn)。

© 2024 NVIDIA Corporation. 保留所有权利。NVIDIA、NVIDIA 徽标、Clara、CUDA、DGX、DGX SuperPOD、Index 和 Triton 均为 NVIDIA Corporation 在美国和其他国家 / 地区的商标和 / 或注册商标。其他公司和产品名称可能是其各自关联公司的商标。其他所有商标均为其各自所有者的财产。3198376。2024 年 3 月 Metadata: please fill in:

