



使用大语言模型 (LLM) 构建 AI 智能体

课程概述

过去几年中，AI 驱动的智能体能力不断提高，新技术不仅使它们能够进行对话，还能使用工具、开展研究，并在大规模范围内执行复杂目标。本课程使您能够开发复杂的智能体系统，这些系统可以执行深度思考、研究、软件调用和分布式操作。在整个课程中，您将获得设计智能体的实践经验，这些智能体能高效地检索和完善信息、智能地路由查询，并利用诸如 LangGraph 等编排工具以及健全的软件工程实践来并发地执行任务。课程结束时，您将建立坚实的智能体架构基础，并能构建有趣的智能体集成来补充现有的工作流和软件栈。



学习目标

参加本次培训，您将学到：

- 理解 LLM 的优势和局限性，以及为什么基于智能体的范式能在现代软件生态中增强其能力
- 学习如何生成结构化输出，以支持机器可解析的函数调用或 API 集成
- 探索检索机制和知识图谱，以获取领域知识
- 使用 LangGraph 等框架，尝试多智能体协作调度
- 实现具有弹性的系统和数据飞轮 (flywheels)，以支持面向生产环境的部署

课程概要

课程时长	8 小时，课后 6 个月内可以继续访问课件和实验 (实验资源用量有限额)
课程模式	讲师授课，每位学员使用完全配置的云端实验环境进行实战练习
预备知识	➤ 深度学习基础知识 (包括注意力机制和 Transformer)。建议具备 DLI 的《深度学习新手入门》或《深度学习基础——理论与实践入门》课程经验。 ➤ 中级 Python 熟练度 (包括面向对象编程及机器学习库的使用)。W3Schools 的 Python 教程或同等实践经验即可满足要求。
使用的工具、库和框架	Python, PyTorch, HuggingFace , Transformers , LangChain , LangGraph
课程测评问题类型	基于技能的编码测试
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，证明在相关领域的技能，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	您需要一台能够上网的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实验环境的专用访问权限
课程价格	微信添加 DLI 小助手 (微信号 NVIDIALearn)，沟通培训需求

课程大纲

课程介绍

(15 分钟)

- > 讲师介绍
- > 登录 DLI 学习平台

智能体抽象与 LLM 基础

(45分钟)

- > 讨论 LLM 的能力与局限性
- > 引入智能体作为任务分解的抽象方式
- > 演示最小化智能体，结合自由文本 LLM 调用

结构化输出与基本执行机制

(45 分钟)

- > 通过基于 JSON 或任务的输出对 LLM 进行瓶颈管理
- > 确保领域对齐与稳定的执行模式
- > 介绍认知架构

休息 (60 分钟)

检索机制与环境工具

(60 分钟)

- > 规范化智能体访问环境的策略，使其能够与其它系统交互
- > 为外部数据存储 (DBs, APIs) 开发工具接口
- > 使用向量 RAG 编码，实现基于文档集的语义检索

知识图谱与文档图谱

(60 分钟)

- > 规划从原始文档到规范化形式的数据演进
- > 设定阈值 / 平衡目标，以驱动事件循环
- > 构建状态池 / 本体 (ontologies)，以实现稳健的领域覆盖

多智能体系统与框架

(60 分钟)

- > 在不同的智能体之间分解任务
- > 规范化通信缓冲区与进程分布方案
- > 区分不同框架及其独特的方法

休息 (15 分钟)

数据飞轮与系统加固

(60 分钟)

- > 捕获使用日志，优化领域约束或子模型
- > 通过人机协作进行纠错
- > 采用真实 / 合成数据进行迭代改进以及工作流简化

扩展与生产部署

(60 分钟)

- > 讨论生产环境相关考虑，如资源管理、并发处理、资源利用率及多租户 (multi-tenancy) 架构
- > 介绍与框架无关的模块化部署 (元框架) 及其选择标准
- > 采用真实 / 合成数据进行迭代改进以及工作流简化

[可选] 实时智能体

- > 讨论多模态交互及与物理世界交互的智能体应用场景
- > 探索机器人、音频系统、世界模型等最新技术进展

[可选] 负责任的智能体

- > 讨论软件设计中的常见失败模式，包括不公平性、法律责任和糟糕的软件体验
- > 研究智能体设计中的制衡系统、标准创建及评估需求

评估测试和总结

(60 分钟)

- > 回顾所学关键内容
- > 在服务器上部署一个智能体断点，实现分布式对话循环，完成评估并获得证书
- > 填写培训调查表

下一步

- 学习更多 DLI 相关课程：
- > [深度学习新手入门](#)
 - > [深度学习基础 —— 理论与实践入门](#)
 - > [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](#)

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。

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)。

© 2025 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。3760505。2025 年 3 月

