

作为开发者，我们如何：

学习大模型

How to become a LLM master



邹雨衡



Datawhale 成员 对外经济贸易大学在读



Logan Zou

邹雨衡



Datawhale 成员 & 对外经济贸易大学在读硕士

个人主页

• <https://logan-zou.github.io/>

Github主页

• <https://github.com/logan-zou>

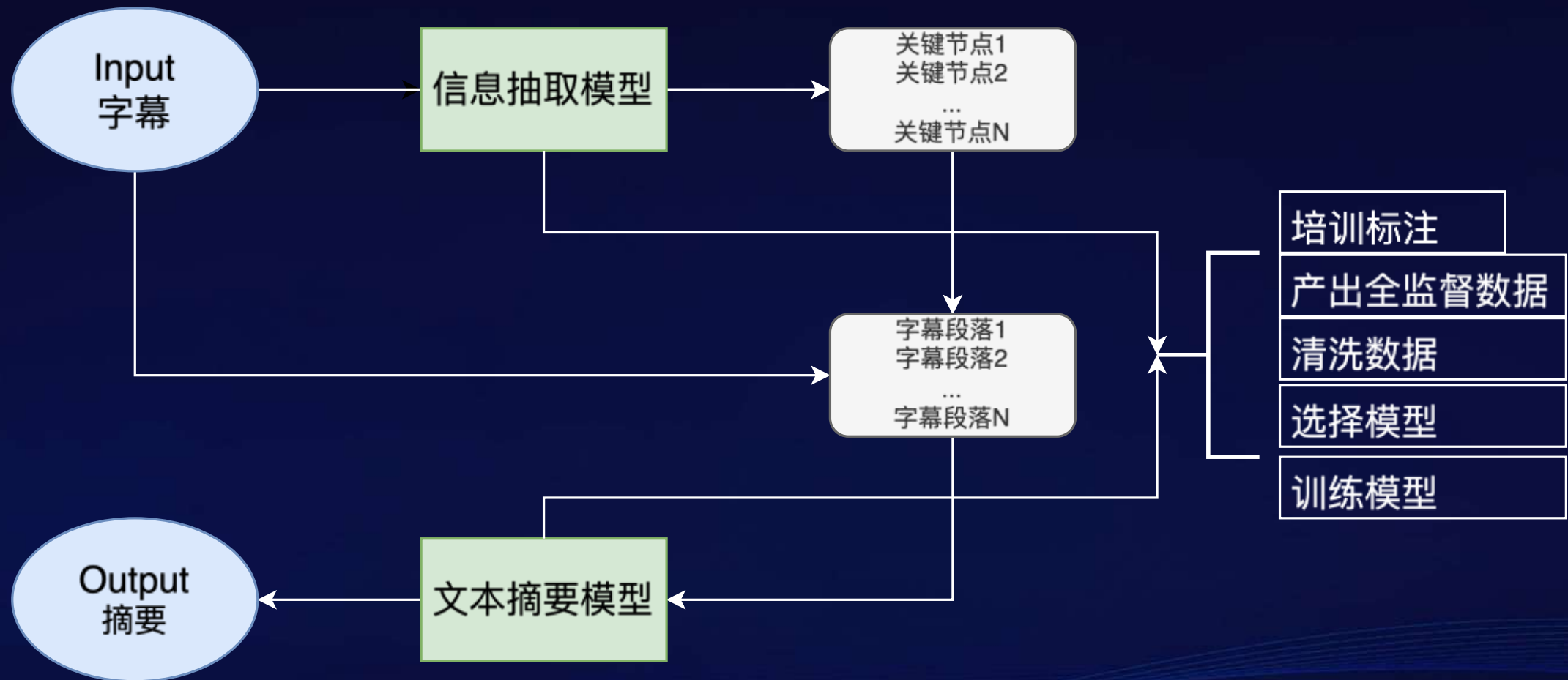
负责项目&成就

- 4个大模型教程负责，10个开源项目核心贡献，6次获得算法赛Top 奖项
- 《LLM Cookbook》面向开发者的大模型手册 Github Star ★ 9.6K
- 《LLM Universe》动手学大模型应用开发 Github Star ★ 2.6K
- 《Self LLM》开源大模型食用指南 Github Star ★ 4.2K

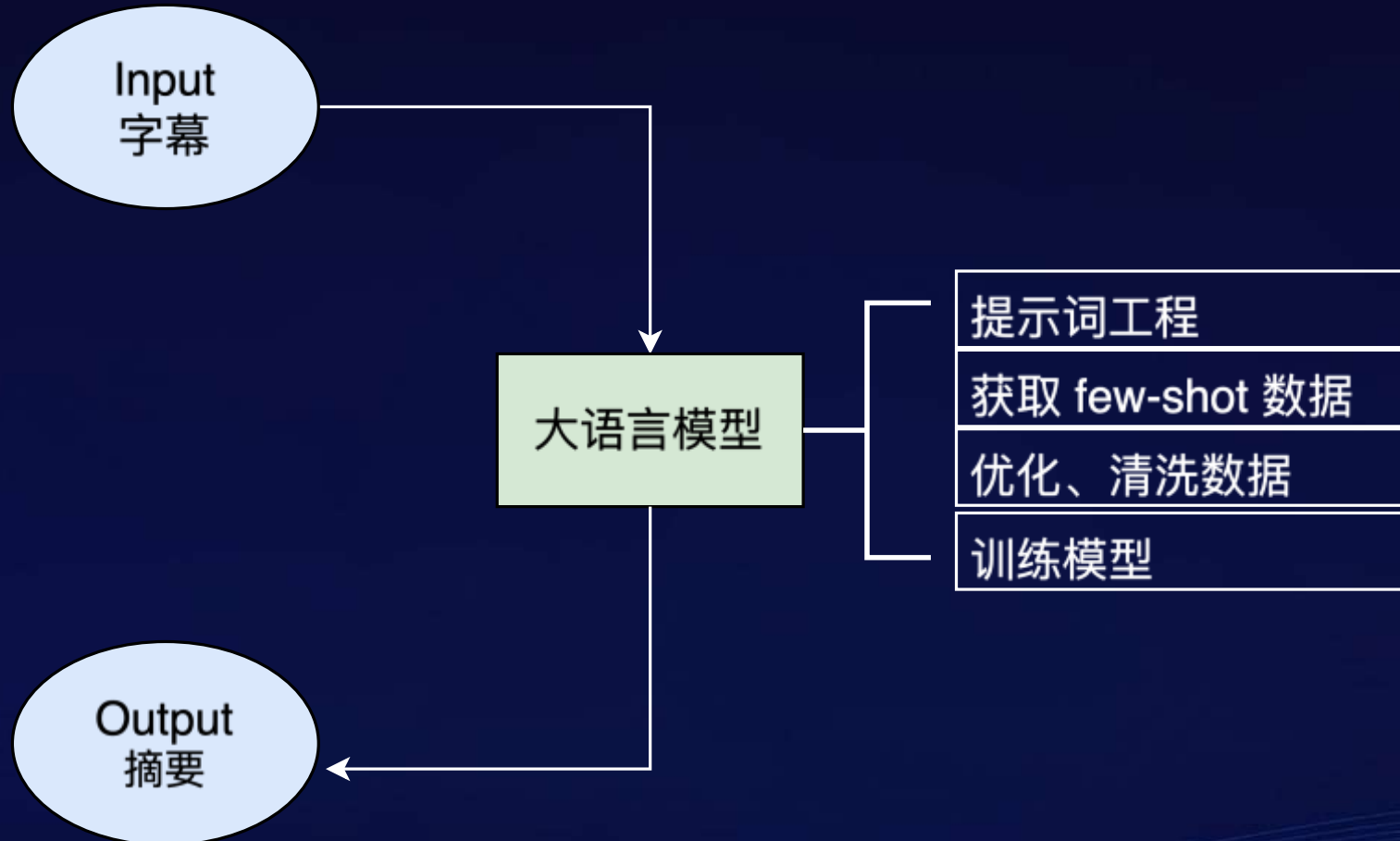
我与 LLM & Datawhale 的一年



传统NLP



LLM



传统NLP → LLM



使用 NLP 模型开发产品的成本与门槛指数级降低



对模型进行训练、结构变动、优化的成本与门槛指数级上升



相较于做模型的“技术活儿”，更重要的是业务与数据的“脏活儿”



更少的数学、代码功底，更多的实践、经验要求——会应用 > 会原理

个人开发者 才是丰富多彩、熠熠生辉的大模型宇宙的未来

如何学习 LLM



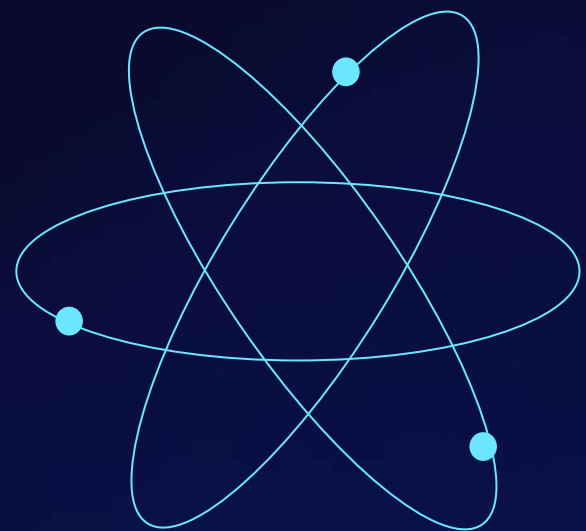
如何学习 LLM



如何学习 LLM 基础



LLM Cookbook——面向开发者的大模型手册



- ▶▶ **面向** • 有足够时间、精力，未确定细分方向，想要全方位入门、学习大模型
- ▶▶ **内容** • 吴恩达大模型课程中文版，充分适配中文语境、中文 LLM
- ▶▶ **课程** • 4门必修课，包括：提示工程、搭建问答系统、LangChain 入门
 - 8门选修课，包括：高级检索、高级 RAG、构建 Agent、微调 LLM、Gradio 等
- ▶▶ **成就** • Github Star 9.6K，视频教程 B 站阅读量 10W+，受人邮出版社邀请出版
- ▶▶ **收获** • 学完本课程，你可以实现大模型全链路、全方位入门
 - 可以为任一方向的深入探究打好基础

课程网址

<https://github.com/datawhalechina/llm-cookbook>

如何学习 LLM 开发



LLM Universe——动手学大模型应用开发

面向对象

- 面向小白开发者
- 基于阿里云服务器，结合个人知识库助手项目，RAG（检索增强生成）的入门教程

主要内容

- 包括大模型简介、调用大模型 API、搭建 RAG应用、验证迭代等

优势所在&达成成就

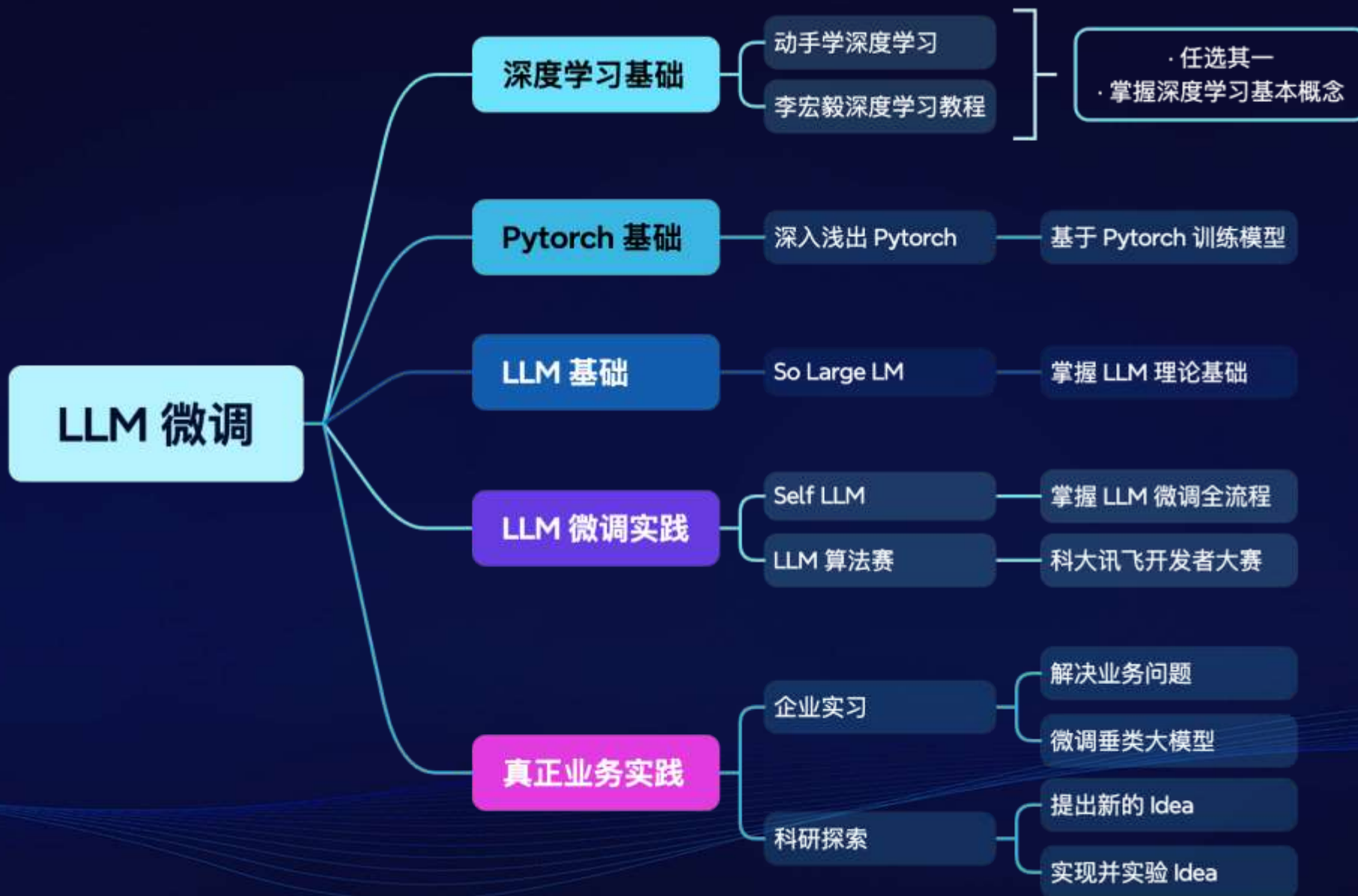
- 最小也最完整的大模型开发教程
- Github Star 2.7K，两次登顶 Github Trending，正在出版流程中
- 用两个小时时间，从算法小白到大模型开发者
- 学完本课程，你将能够动手定制自己的 LLM 应用，掌握 LLM 开发的全部必需能力

网址

<https://github.com/datawhalechina/llm-universe>

LLM Universe
介绍

如何学习 LLM 微调



Self LLM——开源大模型食用指南



Self LLM 介绍

属性&适用范围

- 中国宝宝专属开源大模型教程
- 想要入门大模型微调的算法初学者
- 包括国内外 18个主流开源 LLM
- 调用、部署、微调等全链路手把手教学，不跳过任何一个步骤

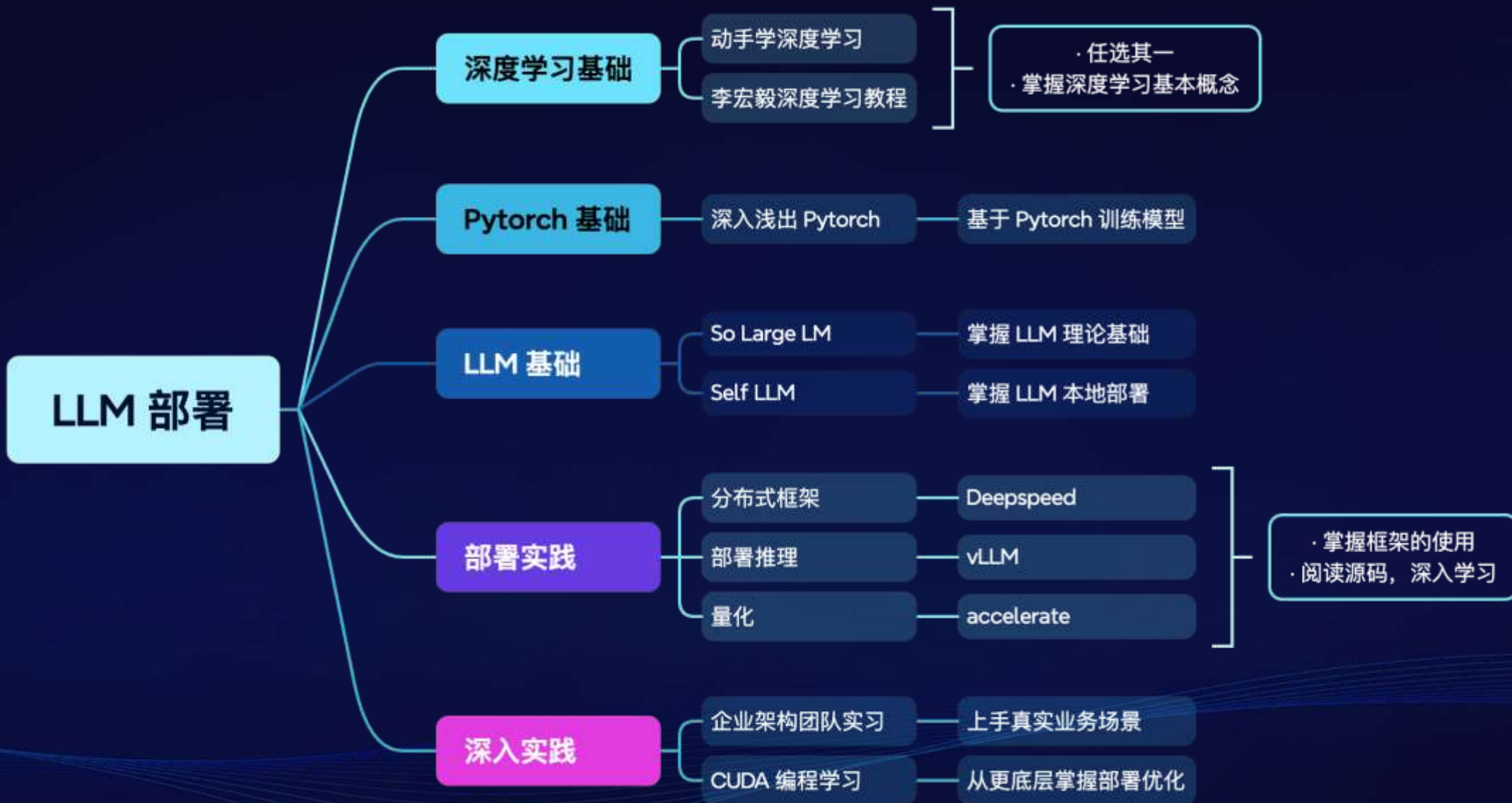
达成成就

- Github Star 4.3K，三次登顶 Github Trending，受人邮出版社邀请出版
- 用一杯奶茶的时间（半天）与成本（12元），速通大模型部署&微调挑战！
- 学完本课程，你能够完全掌握开源 LLM，使用开源 LLM 解决任何实际问题

项目网址

<https://github.com/datawhalechina/self-llm>

如何学习 LLM 部署



如何学习 LLM 训练



LLM 教程链接

- LLM Cookbook, 面向开发者的大模型手册: <https://github.com/datawhalechina/llm-cookbook>
- LLM Universe, 动手学大模型应用开发: <https://github.com/datawhalechina/llm-universe>
- Self LLM, 开源大模型食用指南: <https://github.com/datawhalechina/self-llm>
- Tiny Universe, 大模型白盒子构建指南: <https://github.com/datawhalechina/tiny-universe>
- Hugging LLM, ChatGPT 原理与应用开发: <https://github.com/datawhalechina/hugging-llm>
- So Large LM, 大模型理论基础: <https://github.com/datawhalechina/so-large-lm>
- Hugging Multi-Agent: <https://github.com/datawhalechina/hugging-multi-agent>
- 深入浅出 Pytorch: <https://github.com/datawhalechina/thorough-pytorch>
- 基于 transformers 的 NLP 入门: <https://github.com/datawhalechina/learn-nlp-with-transformers>
- 李宏毅深度学习教程: <https://github.com/datawhalechina/leedl-tutorial>
- 动手学深度学习习题解答: <https://github.com/datawhalechina/d2l-ai-solutions-manual>
- 机器学习公式详解 (南瓜书): <https://github.com/datawhalechina/pumpkin-book>
- 统计学习方法习题解答: <https://github.com/datawhalechina/statistical-learning-method-solutions-manual>

我们希望成为 LLM 与普罗大众的阶梯
以自由、平等的开源精神
拥抱更恢弘而辽阔的 LLM 世界

