

大语言模型

理论、技术与实践

MarkZZZ WeChat: MarkZZZ20XX

课程简介

本课程系统讲授大语言模型（Large Language Models, LLMs）的核心理论与关键技术。课程从自然语言处理的发展历程出发，深入剖析 Transformer 架构的数学原理，全面覆盖预训练、微调、对齐、高效适配、推理部署、提示工程、评估安全等核心主题，并展望多模态、混合专家、长上下文等前沿方向。通过本课程的学习，学员将掌握理解、使用和改进大语言模型所需的理论基础与实践能力。

适合人群

- 计算机科学、人工智能、数据科学相关专业的研究生与高年级本科生
- 从事 NLP、机器学习方向的研究人员与工程师
- 希望深入理解 LLM 原理的 AI 从业者与技术管理者
- 对大模型技术感兴趣的自学者

前置知识

- **数学基础**：线性代数、概率论与数理统计、微积分、最优化基础
- **编程能力**：熟悉 Python 编程，了解 PyTorch 或 TensorFlow 框架
- **机器学习**：了解监督学习、损失函数、梯度下降等基本概念
- **深度学习**：了解神经网络、反向传播、常见网络结构（CNN/RNN）

1 课程内容

讲次	主题	内容概要
1	大语言模型导论	NLP 发展历程、语言模型基础、从 n-gram 到神经语言模型、LLM 的涌现能力
2	Transformer 架构	自注意力机制数学推导、多头注意力、位置编码（正弦/RoPE）、Layer Normalization
3	预训练范式	自回归（GPT）vs 掩码（BERT）、预训练目标函数、缩放定律（Scaling Laws）
4	训练工程	分布式训练（数据并行、张量并行、流水线并行）、混合精度、大规模训练基础设施

讲次	主题	内容概要
5	微调与对齐	监督微调 (SFT)、RLHF 原理、PPO 算法、DPO、Constitutional AI
6	高效微调方法	LoRA、QLoRA、Adapter、Prefix Tuning、参数高效微调的数学原理
7	推理与部署	KV Cache、量化 (INT8/INT4)、推测解码、vLLM、模型蒸馏
8	Prompt 工程与上下文学习	In-context Learning、Chain-of-Thought、RAG、Agent 框架
9	评估与安全	基准测试、幻觉问题、对齐税、红队测试、安全机制
10	前沿与展望	多模态大模型、MoE 架构、长上下文、推理模型、开源生态

2 参考书目

1. Jurafsky, D. & Martin, J.H. *Speech and Language Processing* (3rd ed. draft), 2024.
2. Vaswani, A. et al. "Attention Is All You Need," *NeurIPS*, 2017.
3. Zhao, W.X. et al. "A Survey of Large Language Models," *arXiv:2303.18223*, 2023.
4. 邱锡鹏. *神经网络与深度学习*. 机械工业出版社, 2020.
5. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning*. MIT Press, 2016.