

高校计算机专业大语言模型原理与应用类课程建设研究*

邵非

华东师范大学计算机科学与技术学院, 上海, 200062

摘要 随着生成式人工智能技术的快速发展, 大语言模型 (LLM) 已成为高校计算机专业教学改革的重要方向。本文以《大语言模型原理与实践初步》课程建设为例, 系统梳理当前国内高校同类课程的教学实践与研究现状, 归纳出理论深度与课时冲突、算力门槛高、实验体系失衡、教学内容滞后四大核心矛盾。在此基础上, 提出以轻量化开源模型 MiniMind 为核心载体的课程优化建设方案, 围绕理论精简讲授、低成本实践环境搭建、三层阶梯式实验设计、模块化动态教学大纲四方面展开, 旨在构建兼顾底层原理理解与工程实践能力培养的教学体系。该方案为算力资源有限的本科院校提供了一条可复制、低成本的课程建设路径, 同时融入人工智能伦理与思政教育, 强调技术能力与社会责任并重, 为 AI 前沿课程改革提供实践参考。

关键词 大语言模型, 课程建设, MiniMind, 轻量化模型, 实验教学, 人工智能教育

Research on Course Construction for Large Language Model Principles and Applications in University Computer Science Programs*

Shao Fei

School of Computer Science and Technology
East China Normal University
Shanghai, 200062, China
fshao@cs.ecnu.edu.cn

Abstract—With the swift advancement of generative artificial intelligence technology, Large Language Models (LLMs) have emerged as a pivotal focus for teaching reform in university computer science programs. Taking the development of the course "Introduction to the Principles and Practice of Large Language Models" as an illustrative example, this paper conducts a comprehensive review of the current teaching practices and research landscape for similar courses in Chinese universities, identifying four key challenges: the conflict between theoretical depth and limited class hours, the high computational power requirements, an imbalanced experimental framework, and outdated teaching materials. Building on these insights, we propose an optimized course construction plan centered around the lightweight open-source model MiniMind. This plan encompasses four key strategies: streamlined theoretical instruction, the establishment of a cost-effective practical environment, a three-tiered experimental design, and a modular, dynamic syllabus. Our objective is to create a teaching system that fosters both a deep understanding of foundational principles and the development of practical engineering skills. This approach offers a replicable, low-cost pathway for curriculum development at undergraduate institutions with constrained computational resources. Furthermore, it integrates artificial intelligence ethics and ideological and political education, underscoring the importance of balancing technical proficiency with social responsibility, and provides practical insights for the reform of cutting-edge AI curricula.

Keywords—large language model, curriculum construction, MiniMind, lightweight model, experimental teaching, artificial intelligence education

1 引言

伴随生成式人工智能 (AIGC) 技术跨越式发展, 大语言模型 (LLM) 已然成为新一代人工智能计算体系

的核心基座, 全面重塑计算机学科人才培养逻辑与课堂教学范式, 促使国内各大高校计算机院系陆续开设大语言模型原理与应用相关前沿课程。该类课程旨在打通底层算法理论、全链路工程实践、多行业场景应用三大学习维度, 帮助学生构建系统化大语言模型知识框架。

***基金资助**: 本文得到华东师范大学计算机科学与技术学院 2026 年度研究生课程建设基金资助。

但作为诞生时间较短的新兴课程,现阶段国内高校在教学内容编排、分层实验体系搭建、综合考核评价机制等方面仍处于探索磨合阶段。基于此,本文系统梳理当前国内同类课程教学实践与研究现状,归纳课程落地过程中存在的核心矛盾与现实困境,依托本校《大语言模型原理与实践初步》选修课建设实践,围绕轻量化开源模型 MiniMind 提出适配性建设路径,同时对该类课程长期发展方向作出展望。

2 课程定位与目标体系

从国内高校现有开课方案来看,大语言模型原理与应用课程多设置为计算机科学与技术、人工智能、软件工程等专业高年级前沿选修课,整体目标可划分为知识、能力、素养三层体系。

知识维度:要求学生扎实掌握 Transformer 架构、自注意力机制、预训练范式、多类微调算法等核心底层原理;

能力维度:聚焦模型训练、轻量化微调、模型部署、检索增强生成(RAG)、大模型智能体开发等全流程工程实操能力培育;

素养维度:植入人工智能伦理规范,树立科技自立自强使命感,建立安全合规的 AI 开发思维。

以中国科学院大学同类课程为例,其教学目标明确提出,学生需吃透大语言模型核心思想与关键技术,明晰前沿研究赛道与产业落地路径,依托大模型工具养成系统化分析、解决复杂工程问题的专业能力。该课程明确前置修读要求,需学生提前掌握概率论与数理统计、深度学习、自然语言处理基础课程内容,为本课程理论与实践学习筑牢底层基础[1]。

从教育变革视角分析,计算机专业教学逻辑正从“教会人类编写机器指令”转向“教会人类驾驭人工智能”,教学重心由以代码语法训练为核心,逐步向以概念理解、逻辑推演为核心完成迁移,这一转型特征在国内多所高校前沿 AI 课程改革实践中得到印证[2][3]。

3 核心教学内容与知识体系

综合多所双一流高校公开课程大纲,本课程采用模块化知识架构,四大核心模块覆盖理论、训练、应用、伦理全维度内容:

3.1 基础理论模块

梳理语言模型技术演进脉络,系统讲解传统统计语言模型至 Transformer 架构的迭代逻辑;拆解自注意力机制数学推导、位置编码、层归一化等核心组件运行逻辑;补充预训练数据集采集、清洗、配比等工程基础内容,搭建完整理论认知底座[1]。

3.2 模型训练与优化模块

为本课程重难点章节,内容涵盖大规模预训练基础配置、3D 并行等分布式高效训练方案、显存压缩优化策略、模型参数量与算力消耗估算方法;重点讲解 LoRA 等轻量化微调技术原理与实操流程,兼顾理论推导与工程落地视角[1]。

3.3 应用技术模块

贴合产业真实开发需求,覆盖提示工程(上下文学习、思维链 Prompt)、检索增强生成、大模型工具调用、自主智能体开发四大核心应用技术,是学生学习兴趣最高、实践产出最丰富的内容板块[1]。

3.4 伦理与对齐模块

伴随大模型安全风险持续受到各界关注,对齐学习与 AI 伦理治理成为课程必修内容。课程系统讲解指令微调、人类反馈强化学习(RLHF)等对齐技术;同步融入数据隐私保护、算法偏见治理、生成内容合规等思政教学内容,实现技术教学与价值引领融合。

4 当前主流教学方法与实验设计

针对本课程理论抽象、实操性强的双重特征,国内高校探索形成三类特色教学范式:

4.1 “边讲边码”+项目驱动教学法

采用理论授课、实验实训双线并行模式:课堂同步演示代码运行逻辑,实时拆解模型运行流程;实验课组织学生完整复现训练、微调、部署全流程。期末考试采用小组项目驱动教学法(PBL),3-4 人组队完成大模型应用原型开发,以完整工程作品作为综合评分依据。

4.2 “人机协同”新型教学范式

上海交通大学《智能计算编程基础》提出“人类划定知识框架+AI 辅助代码实现”协同教学模式,教学重点从“训练学生编写基础代码”转向“训练学生精准提问、校验 AI 生成内容”,重点培育人机协同开发能力。该教学模式已在跨学科实训中落地验证,医疗、金融等非计算机背景学生可依托 AI 工具,在 8 周周期内完成需求调研、系统开发、成果演示完整流程。

4.3 概念导向式教学转型

一项面向自然语言处理课程的教改研究提出 LLM 辅助下的 Vibe Coding 教学模式:学生借助 AI 生成代码,同时配套反思式思考题开展过程考核,将评价核心从代码语法流畅度转向底层概念理解。学生教学反馈表明,该模式大幅降低代码调试带来的认知负担,引导学生聚焦注意力于 NLP 核心原理探究。

5 课程建设现存核心难点与内在矛盾

结合一线授课教师教学经验与全国多校公开课程方案,当前大语言模型选修课集中存在四大结构性矛盾:

5.1 理论深度与有限课时的冲突

本课程兼具厚重数学理论与繁杂工程技术栈,课时资源有限是首要教学难题。从教师视角而言,若完整推导 Transformer、自注意力、RLHF 全套数学公式,易造成学生畏难厌学;若简化理论仅讲解 API 调用,则课程沦为浅层工具实训,丧失专业课程深度。从学生基础来看,多数本科生未系统学习分布式并行计算相关知识,讲解大规模模型训练优化内容时难以建立认知衔接。

5.2 实验开展硬件算力门槛偏高

算力资源短缺是制约课程实践落地的现实瓶颈。7B 及以上参数模型的预训练、全参数微调、推理任务均需要大容量 GPU 支撑,普通学生个人笔记本硬件无法满足运行要求。现阶段主流解决方案分为云平台租赁、轻量化量化训练两类:云平台持续产生经费消耗,加重院系教学成本;仅开展轻量化微调实验又让学生无法感知大模型完整训练的算力压力,实践体验存在明显短板。

5.3 实验体系难以平衡原理验证与产业落地

实验设计存在天然取舍难题:若以从零手写 Transformer 作为实验任务,学生能够吃透底层原理,但与工业界成熟开发流程脱节;若直接调用商用 API 搭建 RAG 系统,学生实操门槛低、兴趣浓厚,但易形成只会“调包”、不懂底层的知识盲区。如何搭建分层递进实验体系,兼顾底层原理验证、轻量化模型训练、上层应用开发,适配不同基础学生学习需求,是课程设计的核心考验。

5.4 技术高速迭代与固化教学大纲脱节

大语言模型领域技术迭代速度极快,GPT、Llama、Qwen 系列模型架构、MoE 混合专家、超长上下文等新技术持续更新迭代。固定化教学大纲极易出现内容滞后,学生课下接触前沿产业资讯,课上所学内容存在代差,直接削弱课程教学可信度。如何在稳定整体知识框架的前提下动态补充前沿拓展内容,是课程长期建设需要持续解决的问题。

6 基于 MiniMind 轻量模型的课程优化建设方案

斯坦福 CS224N[3]、清华大学《大语言模型》等标杆课程普遍采用“抓核心、做减法”建设思路:理论

层面深耕 Attention、Transformer 两大核心架构,应用层面聚焦 RAG、智能体两大主流落地场景;实验设计遵循“开源模型推理→轻量化 LoRA 微调→综合性应用项目”递进路径,教学成效得到广泛验证。

笔者所在学校计算机学院拟开设《大语言模型原理与实践初步》选修课,课程建设过程中同样遭遇上述四类教学矛盾。本文以 MiniMind 轻量级开源大模型[4]为核心教学载体,采用理论讲授、分层实验并行建设思路,遵循“核心原理精讲→轻量模型实操训练→预训练与微调全流程复现→前沿技术拓展→综合项目考核”迭代建设路径,针对性化解课程现存痛点,具体优化方案分为四部分:

6.1 理论讲授:抓大放小,聚焦核心架构

本课程定位为专业前沿选修课,不追求知识点面面俱到,优先攻克核心理论难点。理论教学聚焦自注意力、Transformer 架构两大核心模块,其余衍生变体、细分优化技术仅作拓展介绍;以 MiniMind 开源代码作为配套教学素材,推行“源码即教材”教学模式,所有理论知识点均对应模型源码模块进行拆解讲解。

MiniMind 具备极致轻量化特征,最小参数量仅 26M,规模约为 GPT-3 的七千分之一,完整复刻 RMSNorm、RoPE 位置编码、GQA 分组注意力等现代大模型核心组件,无架构简化[4]。学生可单节课完整拆解全套模型代码,无需被千亿级参数的巨型模型抬高学习门槛[6]。

同时社区配套完备的解剖式教学资源:交互式 Jupyter Notebook 将 Tokenizer、模型主体、预训练、监督微调拆分为独立可运行代码块,学生逐段调试即可直观观测张量维度、参数变化[8];配套 24 节系统教程、190 道行业面试题、可视化图解资料,构建零基础到工程落地的完整学习链路,大幅降低学生理论理解门槛。依托上述资源,可有效缓解课时有限、理论晦涩的教学矛盾。

6.2 实践环境:依托轻量模型降低算力门槛

MiniMind 原生适配普通消费级 GPU 开展全流程训练,从根源解决实验算力不足的现实难题。实测环境下,单张 RTX 3090 (24GB 显存) 仅需 2~3 小时即可完整跑完预训练、监督微调全流程,云租赁折算成本仅 3 元左右;笔者使用 RTX 1080Ti 旧款中端显卡,24 小时内也可完成上述全套流程[4][5],详细算力成本/教学方案对比请参见表 1。

基于 MiniMind 的教学方案无需院系投入大额经费采购高端算力集群,学生既可自主租用低成本云主机,也可依托个人中端显卡完成课程实验,全员具备完整复现大模型训练流程的条件,获得“数据→预训练→SFT→LoRA→DPO→推理部署”全流程工程实践体验,具备明显的投入产出优势。

表 1 算力成本/教学方案对比表

教学方案	所需 GPU(显存)	单次实验时间	是否涵盖完整技术流程	预估成本(购置/租用)	适合院校
7B模型全参微调	A100 (40GB) × 4	数天	否	24万元/数千元	顶尖科研院所
7B模型LoRA微调	A100 (40GB) × 1	数小时	否	6万元/数百元	双一流高校
MiniMind全流程	RTX 3090 (24GB) × 1	2-3小时	是	1.2万元/约3元	所有本科院校
MiniMind全流程	RTX 1080Ti (11GB) × 1	24小时	是	0.3万元/自备设备	所有本科院校

6.3 分层实验体系：兼顾底层原理与上层应用

MiniMind 完整覆盖数据预处理、模型搭建、预训练、各类微调、推理部署全链路，为搭建“底层原理-中层训练-上层开发”阶梯式实验体系提供完整素材，三层实验梯度设计内容如下：

(1) 底层原理验证实验

学生自主修改模型维度、网络层数等基础参数，对比观测训练损失变化，直观验证缩放定律(Scaling Law)；调试推理代码，跟踪 KV-Cache 缓存运行逻辑，理解生成加速底层原理，完成抽象理论可视化验证[7]。

(2) 中层核心训练对比实验

开源代码完整支持通用预训练、监督微调(SFT)、LoRA 轻量化微调、DPO 直接偏好优化等训练范式。设置对照实验任务，学生分别运行词语接龙预训练、对话助手监督微调两套脚本，直观区分预训练与微调的目标差异，理解对齐学习核心逻辑。

(3) 上层综合应用开发实验

项目配套 Web 可视化界面、兼容 OpenAI 标准 API 接口，学生完成模型训练后，可快速搭建本地对话机器人，也可对接 FastGPT 等开源工具构建专属知识库应用，打通理论训练到产业落地的完整链路。

上述三层阶梯设计实验充分体现“循序渐进、因材施教”的教学设计理念，灵活适配不同学习层次学生的提升需求，从根本上提升学生基础能力。三层阶梯式实验体系设计架构见图 1。

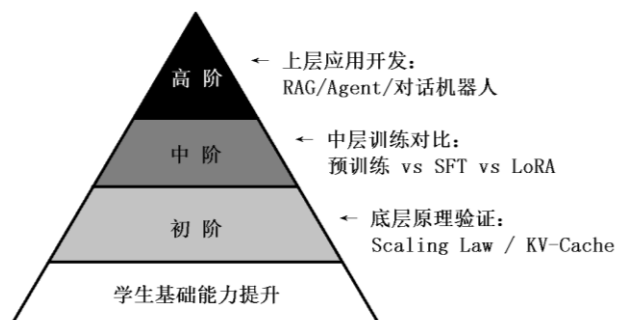


图 1 三层阶梯式实验体系设计架构图

6.4 教学大纲：模块化预留前沿拓展空间

以 MiniMind 为载体搭建的课程框架并非静态固化体系，可随行业技术迭代持续更新教学内容，保障课程前沿性：

(1) 模型架构同步迭代

项目现已更新至 MiniMind-3 版本，底层架构对齐通义千问 3、MoE 混合专家等主流工业模型[5]，学生学习的架构逻辑与当下商用大模型完全统一，不存在知识过时问题。

(2) 配套前沿技术模块。

项目内置 MoE、知识蒸馏、PPO/GRPO 强化学习、工具调用、智能体强化等最新技术完整代码实现[4][5]。行业出现新技术热点时，教师可直接引用项目配套代码补充课堂案例，无需整体重构教学大纲，持续维持课程内容与产业前沿同步。

课程建设方案调研显示，MiniMind 这类轻量化开源模型相当于可拆解的“教学标本”，能够在保证教学质量的基础上将抽象数学原理具象化、巨型模型训练平民化，同时赋予教学大纲持续迭代拓展的空间，打破了传统课程“重推导轻实践”或“重工具轻原理”的两极化困境，实现了底层原理理解与工程实践能力的同步提升，全方位化解课程建设四大核心矛盾。后续将持续根据学生反馈迭代实验任务，并联动本地科技企业补充产业真实案例，进一步完善育人成效

7 结束语

当前生成式人工智能技术仍处于高速迭代周期，持续渗透各行各业，面向计算机专业本科生开设《大语言模型原理与实践初步》选修课，是高校顺应数字产业发展、完善人工智能分层人才培养体系的必然举措。现阶段同类课程普遍存在理论讲解与课时不匹配、实验算力成本高、原理教学与应用实训失衡、教学内容易滞后行业前沿四大现实痛点，而本文提出的以 MiniMind 轻量化开源大模型为核心载体的课程建设方案，为算力资源有限的地方普通本科院校提供一套低成本、易落地、覆盖全学习链路的标准化建设路径。

本文提出四大配套建设策略：理论教学抓核心、精简冗余知识点；依托轻量模型搭建低成本实践环境；设计三层阶梯式实验平衡原理与应用；模块化大纲预留前沿拓展接口。整套方案打破传统课程“重推导轻实践”或“重工具轻原理”的两极化困境。依托MiniMind“源码配套教学”的特色优势，学生可摆脱巨型模型带来的算力与理解双重壁垒，从底层张量运算、注意力机制逻辑起步，完整参与数据集处理、模型预训练、多范式微调、推理部署全流程实操，真正做到吃透底层原理、掌握工程开发能力。整套教学方案同步兼顾理论深度与产业实操训练，同步培育学生专业理论功底与工程落地素养。

技术教学之外，课程全程融入人工智能伦理、数据隐私、算法公平、科技自立自强等思政内容，引导学生树立规范、审慎的AI开发价值观，规避纯粹技术至上的片面认知，培育兼具扎实专业能力与社会责任感的复合型计算机人才。

面向长远发展，大语言模型技术仍将持续演进，多模态融合、超长上下文、自主智能体、新一代MoE架构等新技术会不断涌现。依托轻量化开源模型搭建模块化课程框架，能够持续跟进学术与产业前沿动态，及时更新教学案例，避免课程内容快速老化。该套建设思路具备较强通用性，可迁移至人工智能、软件工程、大数据等多个相关专业，形成可复制、低成本的AI前沿课程改革范式。

总而言之，以MiniMind为代表的轻量化开源模型，有效破解本科阶段大语言模型课程落地的多重阻碍。本文构建的理论精简、算力降本、分层实验、动态更新一体化建设方案，能够系统性化解当前课程建设各类矛盾，为国内高校同类AI前沿课程改革提供可参考

的实践样本。在后续持续教学实践中，还将根据学生课堂反馈迭代优化实验任务，联动本地科技企业补充产业真实项目案例，完善人机协同课堂教学模式，持续提升课程整体育人质量，为国内人工智能产业输送兼具底层理论认知与全链路工程实操能力的后备技术人才。

参考文献

- [1] 中国科学院大学. 大语言模型原理与技术[EB/OL]. (2026-05-22) [2026-08-02]. <https://jwbaucas.ac.cn/sc/course/courseplan/280486>.
- [2] 上海交通大学. 人工智能赋能教育的探索与实践[EB/OL]. (2026-05-22) [2026-08-02]. https://news.sjtu.edu.cn/ztzl_aijyzl/20260522/223014.
- [3] Teaching NLP. CS224n: Natural Language Processing with Deep Learning[EB/OL]. (2026) [2026-08-02]. <https://aclanthology.org/2026.teachingnlp-1.3/>
- [4] jingyaogong. MiniMind[EB/OL]. GitHub, (2025) [2026-08-02]. <https://github.com/jingyaogong/minimind>.
- [5] jingyaogong. MiniMind-3[EB/OL]. Hugging Face, (2026) [2026-08-02]. <https://huggingface.co/jingyaogong/minimind-3>.
- [6] zbohan. MiniMind 极简大模型学习[EB/OL]. 博客园, (2025-11-01) [2026-08-02]. <https://www.cnblogs.com/zbohan/p/19102925>.
- [7] jingyaogong. MiniMind Issues #75[EB/OL]. GitHub, (2025) [2026-08-02]. <https://github.com/jingyaogong/minimind/issues/75>.
- [8] jingyaogong. MiniMind Issues #213[EB/OL]. GitHub, (2025) [2026-08-02]. <https://github.com/jingyaogong/minimind/issues/213>.