

裴晨希



21岁 15227737980 籍贯: 河北省
202218093001@cuc.edu.cn
北京市朝阳区定福庄东街1号
<https://peichenxi77.github.io/midwork>

教育背景

中国传媒大学计算广告双学位在读
2022.09-至今

- 专业方向: 数据科学与大数据技术、广告学专业
- 成绩: GPA 3.89/4.0 93/100
- 排名: 2/39 (前5%)
- 英语水平: CET-6 535 CET-4 553
- 核心课程: 数据结构 (95)、高等数学 (92)、线性代数 (91)、数据库原理与应用 (94)、自然语言处理 (95) 等
- 主要荣誉: 国家奖学金、校三好学生、优秀团干部

竞赛经历

- APMCM 亚太地区大学生数学建模竞赛三等奖 在洪水概率预测问题中, 通过特征处理与梯度提升决策树建模使预测准确率达到95%。
- “互联网+”2024中国国际大学生创新大赛北京市二等奖 在广告投放项目负责投放技术架构设计。
- 第十六届蓝桥杯程序设计大赛Python赛道北京市三等奖

技能&爱好

掌握python、C++、R等编程语言。
掌握 pytorch 框架, 熟练使用 transformer。
了解dify、coze等AI智能开发工具。
掌握SQL、Atom、Vue相关内容, 有网站开发相关实践。

爱好是基于AIGC实现智能绘画, 作品入围金犒奖。喜欢汉服与传统文化。

项目经历

基于ABM与LLM的网络舆情模拟

2024.10——2025.11 科研项目 | 成员

- 项目背景: 智能体在大模型加持下能够模拟人类交互行为, 可以用于社会舆情、传播行为模式等相关内容做研究。
- 个人工作: ①基于HiSim论文的方法, 配合小组成员改进Casevo系统, 设置社会情景: 利用大模型语言能力做人设构建, 利用传播学理论模拟传播效果, 实现舆情复现; ②通过vue架构做数据可视化大屏;。

基于lora微调的广告文案助手

2025.03——2025.04 课程项目 | 负责人

- 项目背景: 广告主文案产出需求与预算不足的矛盾存在。
- 个人工作: ①本地部署llama模型, 在4070GPU环境下完成QWEN-2.5-7B模型的lora微调, 通过8位量化方法压缩参数提升效率, 并尝试了不同参数策略优化微调结果; ②Alpaca格式数据集构建。

预训练模型与LLM结合的自然语言处理系统设计

2024.07——2025.01 北京大数据先进技术研究院 | 实习

- 项目背景: 关于网上大量的开源非结构化新闻数据中蕴含着丰富的信息, 希望开发系统实现新闻分析到结论的全流程。
- 个人工作: ①利用Scrapy、selenium和NewsCrawl进行5000余条新闻收集; ②利用LLM大模型进行数据标注, 通过半监督学习训练 Bert主题分类模型, 使其具有主题扩展性; ③利用reberta-crf模型构建领域自适应框架实现事件抽取, 同时也利用大模型进行抽取对比分析模型效果。

基于对抗训练的多模态谣言鉴别模型

2025.03——至今 国家级大创 | 负责人

- 项目背景: 社交媒体网络不断产生危害性谣言, 如何正确识别谣言成为一项现实挑战。
- 个人工作: ①使用 BERT-Whitening 技术生成混淆性文本与 Diffusion局部图像篡改生成对抗训练样本, 增强模型鲁棒性; ②通过CLIP模型实现文本-图像特征空间映射, 采用对比学习算法优化模态对齐; ③负责小组内协调分工, 进度把握。

