

张扬

(+86) 13126597773

yizhi2026@foxmail.com

求职意向: Agent Harness团队, 或者DeepSeek内任何有需求且认为本人合适的工程/算法团队



基本信息

性别: 男 年龄: 29岁(1997年) 籍贯: 陕西西安 现居住地: 北京

教育经历

- > 布朗大学 硕士, 计算机科学 美国, 普罗维登斯 2018.09 - 2020.10
- GPA: 3.88/4; 主要课程: Computer Vision, Computational Linguistics; 课题方向: 弱监督机器学习; 指导教师: [Stephen Bach](#)
- > 北京航空航天大学 学士, 软件工程 中国, 北京 2014.09 - 2018.07
- 本科期间多次获得各类奖学金; 曾担任数据结构等核心课程助教; 2018年荣获“北京航空航天大学优秀毕业生”称号; 指导教师: [潘海侠](#)

Agent & Vibe Coding & LLM

- 熟悉LLM、Agent原理与技术沿革, 具备SFT、Prompt Engineering工业界经验; 日常工作\生活高频使用Claude Code, Hermes, OpenClaw, OpenCode, Kilo Code, WorkBuddy等国内外多种Agent工具; 对Agent, Harness, Memory等主题具有高度的研究热情。
- NLP/LLM相关经历**: 2017年在北京大学计算机科学技术研究所(王选所)实习, 期间接触Seq2Seq、Transformer等NLP相关新技术(实习成果部分成为2018~2019年本人网络爬虫专著的内容), 其后在CV领域实习(旷视、地平线), 并在推荐广告领域(腾讯、字节)参加全职工作, 近2~3年均致力于LLM范式在推荐广告领域的推广。在2026年2月Anthropic正式提出Harness概念后, 我对这一领域产生极大兴趣, 坚定认同Agent(and AGI) = Model + Harness的理念。对Agent AI领域的工程和设计问题抱有高度好奇心。
- Selected Projects: **AGPA**: 游戏化Vibe Coding的MCP平台; **VibeAware**: 建模和利用用户隐性信号的Harness研究项目, 进行中尚未开源。核心思路是将用户在Agent AI软件中的各类操作和互动建模为 implicit feedback signal, 经 UAHO-9 label space转化为多维label, 使用户与Agent的collaboration成为可建模\预测的监督变量, 将 harness从一个wrapper层升级为可观测、可调控的 policy层。

工作经历

- > 字节跳动, Tiktok 算法工程师 中国, 北京 / 新加坡 2024.10 - 至今
- 任职于Tiktok Live算法团队, 主要项目经历包括:
 - Live题材推荐精排主模型复杂化**: ROW (国际非美欧地区) 精排主模型特征复杂化 & Dense Scale Up
 - 生成式Foundation Model**: 将生成式 (GPT-like) 范式推广到内容AI全链路。作为几位主要贡献者之一, 从0到1参与并贡献了数据链路、样本工程、序列特征工程、模型结构优化、训练吞吐优化等各个环节, 推动生成式模型作为基础模型服务于精排\粗排等下游。
- > 腾讯, 技术工程事业群 算法工程师 中国, 北京 2020.12 - 2024.09
- 任职于机器学习平台部-广告模型平台中心, 主要项目经历包括:
 - 生成式召回系统**: 推动混元大语言模型在微信广告召回场景的落地, 通过LLM输入用户相关query生成广告推荐结果。从0到1完整参与了方案调研、开源模型实验 (Llama, Qwen等)、数据链路打通、基座选型 (1B~13B Dense\MoE模型)、模型训练功能开发、模型精调 (SFT策略优化、SFT样本筛选、Prompt优化、LoRA等PEFT方法) 各个环节, 推动项目成功落地。
 - 广告模型全链路NAS方案**: 全链路网络结构搜索 (NAS) 在召回、粗排、精排模型中的全面落地, [相关论文](#)。
 - 训练超参数优化(HPO)系统**: 推动了HPO系统在训练平台的落地, 参与算法、调度、平台API等模块设计与开发。其中重点主导了各类优化算法 (Epsilon-Greedy、贝叶斯优化、种群进化等) 的实现和接入。开发创新性的HPO算法“FlexHB”, [相关论文](#);

论文专著

- C Ge, X Wang, Z Zhang, Y Qin, H Chen, H Wu, Y Zhang, Y Yang, W Zhu. *Behavior importance-aware graph neural architecture search for cross-domain recommendation*. In Proceedings of the **AAAI Conference on Artificial Intelligence 2025**.
- H Wei, Y Yang, S Ma, H Wu, Y Tang, M Liu, Y Zhang. *AutoPooling: Automated Pooling Search for Multi-valued Features in Recommendations*. Proceedings of the 17th ACM International Conference on Web Search and Data Mining (**WSDM 2024**).
- Y Zhang, H Wu, Y Yang. *FlexHB: a More Efficient and Flexible Framework for Hyperparameter Optimization*. **arXiv:2402.13641**.
- Y Zhang, M Liu, H Wei, Z Hou, Y Tang, H Wu, Y Yang. *Boost Supernet Training with Contrastive Learning and Convergence-Considering Sampling Method*. **CVPR NAS Competition 2022, First Place Prize**.
- 吕云翔、张扬、韩延刚。《Python网络爬虫从入门到精通》. 机械工业出版社, 2020年
- 吕云翔、张扬。《Python网络爬虫实战》. 清华大学出版社, 2019年

其他信息

- 本科期间工业界实习: **旷视科技**, 指导老师: [Gang Yu](#) (现阶跃星辰首席算法研究员)、**威睿** (VMware虚拟机系统开发) 等。
- 硕士期间工业界实习: 于**地平线机器人**美国硅谷实验室实习, 指导老师: [Xianming Liu](#) (现小鹏自动驾驶负责人)。
- 多次获得腾讯内部“业务突破奖”等, 对司内TensorFlow等框架项目均有开源贡献, 在腾讯就职4年内3次获最高绩效 (共7次绩效)。
- Github: [Yizhi Zhang](#) / X: [Yizhi Zhang](#) / LinkedIn: [Yang Zhang](#)