Zishuo Zhao

Hangzhou, China
+86 17635859833
LLM Agents & Reinforcement Learning
zhaozsh@alumni.sysu.edu.cn
zishuo-zhao.github.io
Google Scholar | GitHub

Education

Sun Yat-sen University *Sep 2021 - Jun 2024*

Master's degree in Intelligent Control

School of Intelligent Systems Engineering, Shenzhen, China

Sun Yat-sen University *Sep 2017 - Jun 2021*

Bachelor's degree in Computational Materials

School of Materials, Guangzhou, China

Work Experience

Alibaba Holding *Aug 3, 2026 - Present*

Large Language Model Algorithm Engineer

- *Code Agent Data Synthesis:* Generate synthetic training data for code agents across Terminal-Bench 3.0, machine learning, and scientific problem-solving tasks.
- *Multimodal Agent Training:* Train multimodal agents with reflection, OPD, tool calling, revisiting prior observations, and reinforcement learning.
- *Research on Agent Learning:* Study Claude's reasoning trajectories through probing and reverse engineering; explore distillation, OPD, RSI, and agentic reinforcement learning.

Alibaba International *Jul 4, 2024 - Jul 3, 2026*

Large Language Model Algorithm Engineer

- *Customer service agents:* Contributed to AliExpress pre-sales and after-sales agents, including planner-reasoner architectures, tool calling, reward evaluation, and alignment with DPO and GRPO.
- *Multimodal review:* Worked on image-based review with multimodal models, OCR, image classification, and preference alignment.
- *Voice agents:* Worked on Lazada inbound and ICBU outbound voice agents, covering planning, knowledge grounding, interruption handling, and order understanding.

Internship Experience

Alibaba International / *Research Intern* *Dec 2023 - Feb 2024*

RLHF/DPO and automated evaluation for customer service LLMs.

Tencent AI Lab / *Research Intern* *May 2023 - Oct 2023*

LLM role-playing, game commentary/reporting, and PPO-based context extension.

Publications

[1] **Zishuo Zhao**, Kai Chen, Ao Li, Yuan Liu.
Turn-level Multiscale Density Ratio Estimation for LLM Agents.
arXiv:2609.16760, 2026.

[2] **Zishuo Zhao**, Ziyang Ma, Zhenzhou Lin, Jingyou Xie, Yinghui Li, Ying Shen.
Source-free Domain Adaptation for Aspect-based Sentiment Analysis.
LREC-COLING 2024, pp. 15076-15086.

[3] **Zishuo Zhao**, Yuexiang Xie, Jingyou Xie, Zhenzhou Lin, Yaliang Li, Ying Shen.
Source-Free Unsupervised Domain Adaptation for Question Answering.
IEEE ICASSP 2023, pp. 1-5.

[4] Zhenzhou Lin, **Zishuo Zhao**, Jingyou Xie, Ying Shen.
Hierarchical Type Enhanced Negative Sampling for Knowledge Graph Embedding.
ACM SIGIR 2023, pp. 2047-2051.

[5] Jingyou Xie, **Zishuo Zhao**, Zhenzhou Lin, Ying Shen.
Multimodal Graph Learning for Cross-Modal Retrieval.
SIAM SDM 2023, pp. 145-153.

[6] Jingyou Xie, Jiayi Kuang, Zhenzhou Lin, Jiarui Ouyang, **Zishuo Zhao**, Ying Shen.
FLEX-CLIP: Feature-Level GEneration Network Enhanced CLIP for X-shot Cross-modal Retrieval.
arXiv:2411.17454, 2024.

Technical Skills

Programming and frameworks: Python, verl.
Methods: DPO, GRPO, RLHF, LLM agent development, and multimodal learning.
Languages: English (CET-6, CET-4).

Research Interests

My research focuses on **large language model post-training and agents**, including OPD, synthetic data generation for code agents, RSI-based self-improvement, and multimodal agents.

- **Agentic reinforcement learning:** Multi-turn preference learning, turn-level credit assignment, and alignment for LLM agents.
- **Natural language processing:** Source-free domain adaptation for question answering and aspect-based sentiment analysis.
- **Multimodal learning:** Cross-modal retrieval, multimodal graph learning, and vision-language applications.