黄泽宇

zeroyuhuang.github.io | Google Scholar | zeroy.huang@gmail.com

教育背景

爱丁堡大学信息学院，博士（自然语言处理） 2027 年 6 月预计毕业
2026 Apple PhD Scholar in AI/ML 英国爱丁堡

北京航空航天大学，电子信息硕士 2023 年 6 月
GPA 3.9/4.0；专业排名 2/86 中国北京

北京航空航天大学，信息与计算科学学士 2020 年 6 月
GPA 3.7/4.0；专业排名 6/87 中国北京

研究方向

我主要关注大语言模型的学习机制，包括：（1）学习范式：模型如何从纠错、示范、奖励和上下文中学习；（2）优化目标：训练目标如何塑造、影响模型行为模式；（3）模型架构：模块化、MoE 中的路由与注意力、残差等模块如何影响模型的训练过程。

研究与实习经历

阿里星研究实习生，Qwen 预训练团队 2026 年 7 月–至今
指导人：Bo Zheng、Zihan Qiu 中国北京

字节跳动 Seed 语音团队，TopSeed 研究实习生 2026 年 3 月–7 月
指导人：Wenhao Zhu、Shanbo Cheng 中国上海

Google DeepMind 持续学习团队，学生研究员 2025 年 8 月–2026 年 2 月
指导人：Marc'Aurelio Ranzato、Adhiguna Kuncoro 英国伦敦

INF Technology，研究实习生 2024 年 3 月–8 月
指导人：Zili Wang 远程

北京智源人工智能研究院（BAAI），研究实习生 2022 年 9 月–2023 年 6 月
指导人：Jie Fu 中国北京

腾讯微信 AI，研究实习生 2021 年 11 月–2022 年 5 月
指导人：Yikang Shen 中国北京

论文发表

* 表示共同第一作者。比例按相应赛道投稿数计算。

Zihan Qiu*, **Zeyu Huang***, Kaiyue Wen*, Peng Jin*, Bo Zheng*, et al.. A Unified View of Attention and Residual Sinks: Outlier-Driven Rescaling is Essential for Transformer Training. **COLM 2026 Oral (24/2,939; 0.82%)**

Zeyu Huang, Tianhao Cheng, Zihan Qiu, Zili Wang, Yinghui Xu, Edoardo M. Ponti, Ivan Titov. Blending Supervised and Reinforcement Fine-Tuning with Prefix Sampling. **ICML 2026**

Yijun Yang*, **Zeyu Huang***, Wenhao Zhu, Zihan Qiu, Fei Yuan, Jeff Z. Pan, Ivan Titov. A Controllable Examination for Long-Context Language Models. **NeurIPS 2025 D&B Spotlight (56/1,995; 2.81%)**

Zihan Qiu*, Zekun Wang*, Bo Zheng*, **Zeyu Huang***, Kaiyue Wen, et al.. Gated Attention for Large Language Models: Non-linearity, Sparsity, and Attention-Sink-Free. **NeurIPS 2025 Best Paper Award (3/21,575; 0.014%)**

Zihan Qiu*, **Zeyu Huang***, Bo Zheng, Kaiyue Wen, Zekun Wang, et al.. Demons in the Detail: On Implementing Load Balancing Loss for Training Specialized Mixture-of-Expert Models. **ACL 2025**

Zeyu Huang, Zihan Qiu, Zili Wang, Edoardo M. Ponti, Ivan Titov. Post-hoc Reward Calibration: A Case Study on Length Bias. **ICLR 2025**

Zihan Qiu*, **Zeyu Huang***, Shuang Cheng, Yizhi Zhou, Zili Wang, Ivan Titov, Jie Fu. Layerwise Recurrent Router for Mixture-of-Experts. **ICLR 2025**

Ka Man Lo*, **Zeyu Huang***, Zihan Qiu*, Zili Wang, Jie Fu. A Closer Look into Mixture-of-Experts in Large Language Models. **NAACL 2025 Findings**

Wenyu Du, Tongxu Luo, Zihan Qiu, **Zeyu Huang**, Yikang Shen, et al.. Stacking Your Transformers: A Closer Look at Model Growth for Efficient LLM Pre-Training. **NeurIPS 2024 Spotlight (326/15,671; 2.08%)**

Wenyu Du, Shuang Cheng, Tongxu Luo, Zihan Qiu, **Zeyu Huang**, et al.. Unlocking Continual Learning Abilities in Language Models. **EMNLP 2024 Findings**

Zihan Qiu*, **Zeyu Huang***, Jie Fu. Unlocking Emergent Modularity in Large Language Models. **NAACL 2024 Outstanding Paper Award (6/2,434; 0.25%)**

Zihan Qiu*, **Zeyu Huang***, Youcheng Huang*, Jie Fu. Empirical Study on Updating Key-Value Memories in Transformer Feed-forward Layers. **ICLR 2024 Tiny Papers**

Zeyu Huang, Yikang Shen, Xiaofeng Zhang, Jie Zhou, Wenge Rong, Zhang Xiong. Transformer-Patcher: One Mistake Worth One Neuron. **ICLR 2023**

Zeyu Huang, Wenge Rong, Xiaofeng Zhang, Yuanxin Ouyang, Chenghua Lin, Zhang Xiong. Token Relation Aware Chinese Named Entity Recognition. **ACM TALLIP**

Zeyu Huang, Xiaofeng Zhang, Jun Bai, Wenge Rong, Yuanxin Ouyang, Zhang Xiong. Solving Math Word Problems Following Logically Consistent Template. **IJCNN 2023**

Xiaofeng Zhang*, Yikang Shen*, **Zeyu Huang**, Jie Zhou, Wenge Rong, Zhang Xiong. Mixture of Attention Heads: Selecting Attention Heads Per Token. **EMNLP 2022**

预印本

Zeyu Huang, Adhiguna Kuncoro, Qixuan Feng, Jiajun Shen, Lucio Dery, Arthur Szlam, Marc'Aurelio Ranzato. Context Training with Active Information Seeking. **arXiv 2026**

Tianhao Cheng*, **Zeyu Huang***, Zihan Qiu, Yu Cheng, Edoardo Ponti, Yinghui Xu, Ivan Titov, Zenglin Xu. The Cancellation Hypothesis in Critic-Free RL: From Outcome Rewards to Token Credits. **arXiv 2026**

荣誉与奖项

- **2026 Apple PhD Scholar in AI/ML**（全球 20 人）
- **NeurIPS 2025 最佳论文奖**（主赛道 3/21,575，0.014%）
- **NAACL 2024 杰出论文奖**（6/2,434，0.25%）
- EPSRC DTA 奖学金（2024–2027）
- 研究生国家奖学金（2021–2022；年级排名 3/291）
- 北京航空航天大学优秀研究生（2020–2021）

其他经历与技能

学生工作：北京航空航天大学中法工程师学院研究生会主席（2020 年 12 月–2021 年 12 月）；Orange Lab 与中法工程师学院联合项目负责人（2020 年 10 月–2021 年 5 月）。

国际交流：摩洛哥卡萨布兰卡中央理工学院（2019 年 11 月）；法国 CentraleSupélec（2018 年 2 月）。

编程与框架：Python、C、Shell、Java、LaTeX、MATLAB；PyTorch、JAX、TensorFlow、Hugging Face、Flax 等。

语言：中文（母语）、英语（雅思 7.5）、法语（B2）。