Yuantao Zhang

Tel: 17850565385 (CHN), 87310540 (SIN)
Email: yuantaozhang@u.nus.edu
yuantaozhang@link.cuhk.edu.cn
Website: https://zyttt-coder.github.io/
Address: 7 Faber Drive, 129338, SIN
Wechat Number: 17850565385

Personal Profile

I achieved top 10% GPA in Bachelor of CS at CUHK(SZ) and near-perfect grades in the MComp (AI) program at NUS. I've engaged in research since sophomore year across **3** teams — CUHK(SZ) FreedomAI, NSCCSZ, and NUS NExT++ — collaborating with distinguished professors including Haizhou Li, Tat-Seng Chua, Baoyuan Wu, and Ming Yan, etc. I've published or co-authored **2** papers (one in ACL Findings 2025 and one on arXiv), and my research fields cover LLM alignment and reasoning, diffusion models, federated optimization. I am also **interested in exploring new fields** such as embodied intelligence and LLM4Sci. I am now **open to both Ph.D. and industry opportunities**!

Education

National University of Singapore 2024.8 – Present
Master of Computing (Artificial Intelligence Specialisation) *GPA: 4.7/5*

Nanyang Technological University, Singapore 2023.1 – 2023.5
Undergraduate Exchange Program

The Chinese University of Hong Kong (Shenzhen) 2020.9 – 2024.7
Degree - Bachelor of Engineering with Honours, First Class *CGPA: 3.706 / 4 (top 10%)*
Major - Computer Science and Engineering *MGPA: 3.798 / 4 (top 10%)*

Papers & Patents

M. Li*, **Y. Zhang***, W. Wang, W. Shi, Z. Liu, F. Feng, T. Chua. (2025). *Self-Improvement Towards Pareto Optimality: Mitigating Preference Conflicts in Multi-Objective Alignment.*
2025 ACL Findings (*Equal contribution). [Arxiv, GitHub]

Y. Zhang, Z. Yang. (2025). *A Perplexity and Menger Curvature-Based Approach for Similarity Evaluation of Large Language Models.*
Arxiv Preprint [Arxiv, GitHub]

B. Yang, Z. Yang, Y. Zhang. (2024). *An Approach for Generating Diverse Images Based on Latent Diffusion Models.* CN Patent No. 202411135613.6.

Research & Internships

Research Intern on LLM Alignment 2024.8 – Present
Mentor: Prof. Wenjie Wang, Prof. Tat-Seng Chua *NExT++ Lab, NUS*

1. **Research on resolving preference conflicts in Multi-Objective Alignment (MOA)**

- **Background**: MOA aims to align users' different preferences across multiple objectives. However, preliminary experiments revealed that DPO-based MOA methods are affected by **preference conflicts**, where an answer's preferences vary across objectives. Preference conflicts impact the alignment **Pareto front** and are difficult to resolve using simple methods.
- **Verification of preference conflicts' negative impact**: Data with different preference conflict proportions (0, 30%, 60%, 90%) were used to align **2 MOA benchmarks** and **2 sets of objectives**. As the conflict proportion increased, the performance decline became more significant. The performance at a 90% conflict proportion was even **similar to that before alignment**.
- **Proposed SIPO**: By constructing **Pareto-optimal** answers that outperform original response pairs on all objectives, we proposed an **automated self-improvement framework** to resolve preference conflicts. The answers were achieved through sampling, self-refinement, and self-filtering of

first-aligned models and used for secondary alignment. Experiments on **LlaMA-2-7B** and **Alpaca-7B** demonstrated the framework's superiority. On **BeaverTails**, it could improve the **helpfulness reward by 2.1** and the **safety reward by 3.0** compared to the best baseline.

2. **Research on generalization of math reasoning to factual QA tasks**

- **Background and preliminary findings**: GRPO endows base models with math reasoning abilities but affects their factuality. This study found that models did not forget knowledge, but **collapse on hidden states**, leading to mis-representation of knowledge. Moreover, **interventions towards the direction of correct knowledge** make factual reasoning performance better (validated by experiments on **Qwen2.5-1.5B**, **Llama3.1-8B**, and **Mistral-v1-7B**). The research target is to modify GRPO with minimal costs to improve factual reasoning while retaining math reasoning.
- **Proposing meta-guided verifiable reward (ongoing)**: Introducing **meta-reward** in GRPO to dynamically scale the verifiable rewards for math tasks. The meta-reward is calculated based on **a linear probe** to guide the hidden state in training towards correct knowledge direction.

Research Intern on LLM similarity evaluation 2024.3 – 2024.7
Mentor: Dr. Zhankui Yang (Postdoc) *National Supercomputing Center in Shenzhen*

- Conceptualized a metric utilizing **Perplexity Curves**, which plot the perplexity of sub-sequences in a sentence against word indices, and **Menger Curvature Difference** of these curves to assess the similarity between LLMs. This metric can aid in **detecting unethical practices in LLM usage**, such as slightly altering an existing LLM to falsely claim a new development. Experiments across **9 LLMs** on **Wikipedia, medical, and legal datasets** show its effectiveness.

Projects

Assist in writing "Overview-of-ChatGPT" 2023.9 – 2023.12
Mentor: Dr. Feng Jiang (Postdoc), Prof. Haizhou Li *Freedom AI Team, CUHK(SZ)*

- Responsible for the "Post-GPT Era" part of "Overview-of-ChatGPT". Summarize the structure, training methods and performance of mainstream open-source LLMs, including GLM, Baichuan, WizardLM, Mistral, etc., as well as the Phoenix and HuatuoGPT models developed by the team. [GitHub]

Develop a data Distillation framework using the Stable Diffusion (SD) model 2023.5 – 2023.8
Mentor: Dr. Haonan Wang, Prof. Kenji Kawaguchi

- **Background**: Dataset distillation aims to synthesize a small dataset to replace the complete dataset and achieve similar training performance. Traditional methods generate distilled data by pixel level optimization or latent space optimization of generative models, but suffer from low generalization.
- Optimized the **token embeddings** of SD to generate distilled images. Embeddings could be further modified to adapt training on various base models. The framework was tested on a **10-class subset of ImageNet**, achieving SOTA results while improving controllability and generalization. [GitHub]

Awards & Honors

NUS: AY2024-25 Semester 1 Dean's List
CUHK(SZ): AY2020-21, 21-22, 22-23, 23-24 SDS Dean's List | AY2022-23 SDS TA/USTF Award | Undergraduate Research Award | 2021 Programming Contest, Third Place (Advanced Group) | AY2021-22 Academic Performance Scholarship: Class C (20000 RMB) | Bowen Scholarship (35000 RMB / Year)

Languages & Skills

Languages – English (**GRE**: 328 + 3.5; **TOEFL**: 102, Reading 27, Listening 28, Speaking 22, Writing 25), Mandarin (native)
Programming Languages – Python, C++, Java, Verilog
Technical tools – Pytorch, VERL, Git, SQL, Django
Organization experience – Teaching Assistant for Data Mining course; Group Leader in the Freshmen Induction Program; Member of the campus entrepreneurship team