Yuxuan TONG

E-Mail: tongyuxuan361@gmail.com | Homepage: tongyx361.github.io

Last Updated on July 15, 2024

EDUCATION

Department of Computer Science and Technology (DCST), Tsinghua University (THU) 2021.09 - Present

Bachelor's degree in progress

GPA: 3.8/4.0

PUBLICATIONS

DART-Math: Difficulty-Aware Rejection Tuning for Mathematical Problem-Solving

Under review

Yuxuan Tong, Xiwen Zhang, Rui Wang, Ruidong Wu, Junxian He

ImageReward: Learning and Evaluating Human Preferences for Text-to-Image Generation

*Accepted by **NeurIPS 2023***

Jiazheng Xu*, Xiao Liu*, Yuchen Wu, **Yuxuan Tong**, Qinkai Li, Ming Ding, Jie Tang, Yuxiao Dong

(*) denotes co-first authors

RESEARCH EXPERIENCE

CX Group @ Language Technologies Institute, Carnegie Mellon University (CMU-LTI) 2024.07 - Present

*Research intern, **advised by** Prof. Chenyan Xiong*

Working on **data for on-line (direct) preference optimization algorithms (e.g. DPO)**.

NLP Group @ Hong Kong University of Science and Technology (HKUST) 2023.07 - 2024.06

*Research intern, **advised by** Prof. Junxian He*

Worked on **synthetic data for mathematical reasoning** since 2023.12, before which I worked on **process supervision**, **reward modeling**, **constrained decoding** for **mathematical reasoning** and **model merging**.

Knowledge Engineer Group (KEG) @ Tsinghua University (THU) 2022.11 - 2023.06

*Research intern, **advised by** Prof. Jie Tang and Prof. Yuxiao Dong*

Worked on **reward modeling** and **RLHF** for **text-to-image generation**

PROJECTS

End-to-End Evaluation Model for Text-to-Image Generation 2023.04 - 2023.11

Admitted and completed as **A-Level (top 25%)** in *Tsinghua University Initiative Scientific Research Program in 2023*

Maintainer of the GitHub Repo. `THUDM/ImageReward` 2023.03 - Present

In addition to regular contribution and maintainment, I **integrated our ImageReward into Stable Diffusion Web UI**

AWARDS

- Tsinghua University *Research Scholarship* 2023
- Tsinghua University *Comprehensive Merit Scholarship* (**top 5% undergraduates**) 2022

ACADEMIC SERVICES

Conference Review

NeurIPS 2024

SKILLS

Programming Languages	Python, C/C++, TypeScript/JavaScript, etc.
ML Libraries	PyTorch, DeepSpeed, HuggingFace, vLLM, etc.
Tools	Git, Shell, SLURM, Linux Utilities, LaTeX, etc.