QIYUAN ZHANG

Ph.D. Student, City University of Hong Kong

 83 Tat Chee Avenue, Kowloon, Hong Kong SAR

 (+852) 51640323   qzhang732-c@my.cityu.edu.hk   G Google Scholar

EDUCATION

- **City University of Hong Kong** — Jan. 2023 – Present
  Ph.D. in Computer Science (Advisor: Prof. Chen Ma) — *Hong Kong*

- **University of Electronic Science and Technology of China** — Sep. 2019 – Jun. 2022
  M.Sc. in Computer Science — *Chengdu, China*

- **University of Electronic Science and Technology of China** — Sep. 2015 – Jun. 2019
  B.Sc. in Computer Science — *Chengdu, China*

RESEARCH EXPERIENCE

- **Test-time Scaling in LLMs** — Feb. 2025 – Present
  Ph.D. Research — *Hong Kong*

  - **What, How, Where, and How Well? A Survey on Test-Time Scaling in Large Language Models**: Introduced a novel and hierarchical taxonomy with 4 orthogonal dimensions for deconstructing existing Test-time Scaling literature and provided hands-on guidelines for beginners.

- **Evaluation for LLMs** — Jun. 2023 – Present
  Ph.D. Research — *Hong Kong*

  - **Collaborative Performance Prediction for LLMs**: Introduced a novel method for predicting LLM performance across different tasks using a multi-model benchmark matrix. Discovered new performance predictors beyond computational scaling laws.
  - **Improving Evaluation via Response-Adapted References**: Developed REVISEVAL, a technique that generates response-adapted references, significantly enhancing evaluation reliability.
  - **Unlocking Test-time Scaling in LLM-as-a-Judge**: Proposed a crowd-based comparative evaluation technique that leverages additional responses to improve LLM-as-a-Judge accuracy.

- **Mathematical Reasoning** — Jul. 2019 – Jul. 2021
  M.Sc. Research (Collaborator: Lei Wang) — *Chengdu, China*

  - **Explainable Mathematical Reasoning in Conversational QA**: Created a dataset combining multi-step conversational QA with interpretable reasoning graphs for improved mathematical reasoning.
  - **Toolkit for Math Word Problem Solving**: Developed an open-source framework that became widely used in the research community for mathematical problem-solving.

PUBLICATIONS

- **Qiyuan Zhang**, Fuyuan Lyu, et al., "**What, How, Where, and How Well? A Survey on Test-Time Scaling in Large Language Models**", ArXiv:2503.24235 (Under Review at COLM 2025).

- **Qiyuan Zhang**, Yufei Wang, et al., "**Crowd Comparative Reasoning: Unlocking Comprehensive Evaluations for LLM-as-a-Judge**", ArXiv:2502.12501 (Under Review at ACL 2025).

- **Qiyuan Zhang**, Yufei Wang et al., "**REVISEVAL: Improving LLM-as-a-Judge via Response-Adapted References**", ICLR (Poster), 2025.

- **Qiyuan Zhang**, Fuyuan Lyu, Xue Liu, Chen Ma, "**Collaborative Performance Prediction for Large Language Models**", EMNLP, 2024.

- **Qiyuan Zhang**, Chen Ma, "**Active Instruction Tuning for Large Language Models with Reference-Free Instruction Selection**", PAKDD, 2024.

- **Qiyuan Zhang**, Lei Wang et al., "**NOAHQA: Numerical Reasoning with Interpretable Graph Question Answering Dataset**", EMNLP (Findings), 2021.

- Yihuai Lan, Lei Wang, **Qiyuan Zhang** et al., "**MWPToolkit: An Open-Source Framework for Deep Learning-Based Math Word Problem Solvers**", AAAI Workshop, 2021.