HUARUI XIE

+8615820736135 | xihri412@g.ucla.edu | LinkedIn.com/Huarui Xie

EDUCATION

University of California, Los Angeles, *Los Angeles, CA* 09/2022 - 06/2024
Master of Applied Statistics and Data Science GPA: 3.88/4.00

University of Illinois at Urbana-Champaign, *Urbana-Champaign, IL* 08/2019 - 05/2022
Bachelor of Science in Statistics (minor in Mathematics) GPA: 3.59/4.00

PUBLICATIONS

Xifeng Yao, Chengyuan Ma, Dongyu Lang, Yinhao Ni, Zhiwei Xu, **Huarui Xie**, Zihao Chen, Guang Shen, Dandan Tu, Yi Bai, Changzheng Zhang, "*SLIM: Subtrajectory-Level Elimination for More Effective Reasoning";* **EMNLP 2025 Findings***;* https://arxiv.org/abs/2508.19502*,* (Findings acceptance rate: 17.35%)

Xifeng Yao, Dongyu Lang, Wu Zhang, Xintong Guo, **Huarui Xie**, Yinhao Ni, Ping Liu, Guang Shen, Yi Bai, Dandan Tu, Changzheng Zhang, "*SCoGen: Scenario-Centric Graph-Based Synthesis of Real-World Code Problems";* https://arxiv.org/abs/2509.14281.

EMPLOYMENT

Huawei Celia Department, *Shanghai, China* 10/2024 - Present
AI Engineer/Researchers

- **Designed scalable data pipelines** for curating high-quality domain-specific text datasets and assessed synthesis quality across dimensions such as generalizability, consistency, and complexity.
- **Developed structured evaluation methods and adapted open-source evaluation frameworks** to benchmark LLM performance on tasks (coding, logical reasoning), ensuring accuracy and reliability of model outputs.
- **Conducted failure analysis** to identify model weaknesses and proposed improvements to enhance robustness.
- **Co-authored two research papers (EMNLP, AAAI)** on innovative data projects and **led comprehensive researches** on industry model capabilities.

Palo Alto City Council (Office of Greg Tanaka), *Palo Alto, CA* 03/2024 - 07/2024
AI/Data Analyst Intern

- Built an **end-to-end agent system** by implementing a chatbot backend with Groq Llama3 API; designed prompt structures and reusable LangChain tools, and integrated with the frontend via LangServe FastAPI.
- **Applied statistical modeling and geospatial analysis** to election-related datasets, identifying key voting patterns and providing actionable insights for outreach strategies.

TCL, *Shanghai, China* 08/2023 - 09/2023
AI Data Intern

- **Engineered and refined prompt strategies** for customer service and smart device control applications; leveraged LLM fine-tuning expertise to improve response accuracy.
- **Created and curated 40+ internal datasets** for quantitative model assessment and acceptance testing.
- Collaborated with algorithm engineers, client and domain experts to **gather critical validation cases and address challenging edge-case scenarios** for demos and user acceptance, improving the overall LLM evaluation pipeline for future deployments.

SKILLS

- **LLM Training:** Dataset construction, fine-tuning workflows
- **LLM Deployment & Inference:** Serving via LangChain, LangServe, FastAPI, prompt optimization
- **LLM Evaluation:** Benchmark design, structured metrics, robustness and failure analysis
- **Programming & Tools:** Python, PyTorch, Hugging Face, Git, Linux, Docker
- **Data & Quality:** Data curation, cleaning, synthetic data validation, statistical testing
- **Foundation Knowledge of LLMs:** Transformer architectures, NLP algorithms, evaluation methodologies

SELECTED PROJECTS

SCoGen: Scenario-Centric Graph-Based Synthesis of Real-World Code Problems 05/2025 - 09/2025

- Curated and structured **knowledge graphs** to produce real-world coding scenarios, enabling systematic synthesis of diverse and challenging problem instances.
- Contributed to the development of the **evaluation framework** to systematically assess correctness, difficulty calibration, and pedagogical effectiveness of generated code problems.
- Analyzed **failure cases** from generated problems, identifying gaps in open-sourced models and self-trained models.

SLIM: Subtrajectory-Level Elimination for More Effective Reasoning 03/2025 - 08/2025

- Prepared large-scale datasets by conducting **deduplication, noise reduction, and data cleaning**, ensuring consistency and reliability for model training and evaluation.
- Designed and executed model **evaluation pipelines**, benchmarking coding tasks under different subtrajectory elimination strategies.
- Collaborated with senior researchers to **analyze empirical results**, identify bottlenecks, and propose refinements to the experimental design.
- **Co-authored the manuscript** for submission to a top-tier AI venue, contributing to literature review, methodology explanation, and results interpretation.

Enhanced Object Detection on Road Objects 01/2024 - 05/2024

Advised by Yingnian Wu, UCLA

- **Collected and preprocessed** open-source on-road objects datasets (11,000+ samples) for training and testing.
- **Applied transfer learning** on three YOLOv8 models (n, s, m) by freezing backbone layers and **fine-tuning hyperparameters** to improve on-road object detection.
- **Assessed model performance** on a test set of 1,900+ samples using standard evaluation metrics.