JIAXI LI

University of Georgia, School of Computing
+1-(470)-263-8931 | e: jx.li@uga.edu / jiaxi.li0216@gmail.com
Homepage: plusnli.github.io

EDUCATION

University of Georgia — Georgia, USA
Ph.D. Student — Aug. 2024 - now

- GPA: 3.75/4
- Research Focus: Large Language Models, Multimodal Large Language Models, Machine Reasoning, Healthcare.

Shandong University — Shandong, China
Bachelor of Science in Computer Science and Technology — Sep. 2020 – Jun. 2024

- GPA: 87.8/100
- Selected awards: Three-time recipient of the Third-class Academic Scholarship Prizes (Nov. 2021, Nov. 2022, Nov. 2023)

SELECTED PUBLICATIONS AND PRE-PRINTS

(*Equal Contribution, †Corresponding Author.)

Mitigating Hallucination Through Theory-Consistent Symmetric Multimodal Preference Optimization
Wenqi Liu, Xuemeng Song†, **Jiaxi Li**, Yinwei Wei, Na Zheng, Jianhua Yin, Liqiang Nie
Published in *The Thirty-Ninth Annual Conference on Neural Information Processing Systems (NeurIPS 2025).*

HELENE: Hessian Layer-wise Clipping and Gradient Annealing for Accelerating Fine-tuning LLM with Zeroth-order Optimization
Huaqin Zhao*, **Jiaxi Li***, Yi Pan, Shizhe Liang, Xiaofeng Yang, Fei Dou, Tianming Liu, Jin Lu
Published in *The 2025 Conference on Empirical Methods in Natural Language Processing Main Conference (EMNLP 2025 Main).*

Automating Expert-Level Medical Reasoning Evaluation of Large Language Models
Shuang Zhou*, Wenya Xie*, **Jiaxi Li***, Zaifu Zhan, …, Yucheng Shi, Ninghao Liu, Zirui Liu, Rui Zhang
Published in *npj Digital Medicine.*

MITS: Enhanced Tree Search Reasoning for LLMs via Pointwise Mutual Information
Jiaxi Li, Yucheng Shi, Jin Lu, Ninghao Liu
Submitted to *The Fourteenth International Conference on Learning Representations (ICLR 2026)*.

Fact or Guesswork? Evaluating Large Language Models' Medical Knowledge with Structured One-Hop Judgments
Jiaxi Li, Yiwei Wang, Kai Zhang, Yujun Cai, Bryan Hooi, Nanyun Peng, Kai-Wei Chang, Jin Lu
Submitted to EACL 2026.

Proximal Federated Learning for Body Mass Index Monitoring using Commodity WiFi
Jiaxi Li, Kiran Davuluri, Khairul Mottakin, Zheng Song, Fei Dou, Jin Lu
Published in *ICASSO Workshop of the 30th Annual International Conference on Mobile Computing and Networking (ICASSO@MobiCom 2024).*

RESEARCH EXPERIENCE

UCLA / UC Merced — Remote
Supervised by Prof. Yiwei Wang — Jul. 2024 – Feb. 2025
Research Focus: Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), Healthcare

- Researched on Retrieval-Augmented Generation (RAG) with LLMs and its application in healthcare.
- Finished a first-author work **Fact or Guesswork? Evaluating Large Language Models' Medical Knowledge with Structured One-Hop Judgments**. We build the MKJ dataset, a factuality dataset for assessing LLMs'

factual knowledge. We study the different aspects (performances, calibration, …) of LLMs based on the MKJ dataset and find several interesting observations.

Hong Kong Baptist University (HKBU), Trustworthy Machine Learning and Reasoning (TMLR) Group Remote
Working with PhD Candidate Zhanke Zhou & Supervised by Prof. Bo Han April. 2024 – Now
Research Focus: Machine Reasoning with Foundation Models

- Focus on LLM/MLLM Reasoning, including Test-time Scaling Compute, RL, and Agentic Tool Learning.
- Researched on LLM Jailbreaking and run experiments for the paper **DeepInception: Hypnotize Large Language Model to Be Jailbreaker (Published in NeurIPS 2024 SafeGenAI Workshop)**.
- Build an interesting reasoning dataset based on theft detective scenarios which tests the limitation of LLMs' logical reasoning abilities, which is now in part of paper **From Passive to Active Reasoning: Can Large Language Models Ask the Right Questions under Incomplete Information? (Published in ICML 2025).**
- Participated in developing python package *reasoning-pro*, a comprehensive package for running reasoning-related models, algorithms and benchmarks (Responsible for the benchmark and evaluation part).

Rensselaer Polytechnic Institute, Data Analytics and Machine Intelligence (DAMI) Lab Remote
Supervised by Prof. Yao Ma Sep. 2023 – Jan. 2024
Research Focus: Graph Neural Networks

- Focused on the topic "When do Graph Neural Networks (GNNs) work on node classification and when not".
- Deepened understanding of the fundamental mechanism behind GNNs and why they can perform well under certain heterophily.
- Wrote a blog (Link to Zhihu) to summarize how homophily and heterophily influence GNNs from a unified perspective based on papers Is Homophily a Necessity for Graph Neural Networks?, Demystifying structural disparity in graph neural networks: Can one size fit all?, and other relevant papers.
- Conduct extensive experiments to explore further research questions (e.g., whether class-wise homophily has relation with class-wise prediction accuracies; to develop evaluation metrics to quantify the discriminative ability of GNN after each aggregation step).

Shandong University, Information Retrieval Lab (IRLab) Qingdao, China
Supervised by Prof. Zhaochun Ren Feb. 2023 – Aug. 2023
Research Focus: Large Language Models

- Participate in training a Retrieval-Augmented LLM for legal judgements, which is open-sources in GitHub (link).
- Innovated new evaluation metrics, utilizing advanced LLMs such as GPT-4, Vicuna, and ChatGLM, specifically tailored for Natural Language Processing tasks like Text Summarization and Open Domain Question Answering.
- Implemented cutting-edge techniques including Prompt learning, Parameter-efficient learning, and Finetuning to fine-tune LLMs for optimal performance.

Work Experience

Shandong Houde Measurement and Control Technology Co., Ltd. Jinan, China
Research Intern, Software and Research Department Jun. 2022 – Sep. 2022

- Assisted the department in the research and development of the corn yield measurement control system and a third-party testing system for an integrated development environment, with a focus on deep learning for data analysis
- Participated in the optimization and deployment of natural language models during the software development process