Liqin Ye

Email: liqiny@gatech.edu | Webpage: liqinye.github.io | LinkedIn: in | Github: ○

EDUCATION

• **Georgia Institute of Technology** — Atlanta, GA, USA
M.S. in Computer Science — *Aug 2023 - Dec 2025*
Ph.D. in Machine Learning — *Aug 2024 - Jun 2029 (Expected)*
- **GPA**: 3.84/4.00
- **Research Interest**: LLM reasoning and agent, LLM alignment to diverse user preferences, LLM in Finance.
- **Advisor**: Dr. Chao Zhang and Dr. Sudheer Chava

• **University of California, Irvine** — Irvine, CA, USA
B.S. in Computer Science (ICS Honor Student) — *Aug 2019 - Jun 2023*
- **GPA**: 3.95/4.00 (Magna Cum Laude)
- **Core Courses**: Deep Generative Models, Deep Learning, Graph Algorithms, Formal Language & Automata
- **Advisor:** Dr. Stephan Mandt

PUBLICATIONS

1. **Liqin Ye**, Agam Shah, Chao Zhang, Sudheer Chava. **Calibrating Pre-trained Language Classifiers on LLM-generated Noisy Labels via Iterative Refinement**. In *Proceedings of the 31st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD**)*, 2025.

2. Rongzhi Zhang*, **Liqin Ye***, ..., Chao Zhang. **Precise Attribute Intensity Control in Large Language Models via Targeted Representation Editing**. Submitted to *ICLR* 2026.

3. Agam Shah, **Liqin Ye**, ..., Wei Xu, Sudheer Chava. **Beyond the Reported Cutoff: Where Large Language Models Fall Short on Financial Knowledge**. In *Proceedings of the 2nd Conference on Language Modeling (**COLM**)*, 2025.

4. Agam Shah, ..., **Liqin Ye**, ..., Sudheer Chava. **Words That Unite The World: A Unified Framework for Deciphering Central Bank Communications Globally**. In *Proceedings of the 39th Annual Conference on Neural Information Processing Systems (**NeurIPS**)*, 2025.

RELEVANT EXPERIENCE

• **Graduate Research and Teaching Assistant** — *May 2024 - Present*
Financial Services Innovation Lab, Georgia Institute of Technology — Atlanta, GA, USA
- Developed a test-time intervention via representation editing for multi-objective alignment, achieving **26.16%** success rate and reducing **8x** computational overhead. Co-first authored paper is submitted to ICLR 2026.
- Built a denoising framework to calibrate pre-trained classifiers under LLM-generated noisy labels via simplex diffusion model, boosting accuracy by **7.3%**. First-authored paper is accepted at KDD 2025.
- Created a **197k+** revenue questions dataset to unveil the temporal and cross-sectional financial knowledge gap in LLMs. Linked hallucination rates with knowledge errors. Co-authored paper is accepted at COLM 2025.
- Curated a world central banks dataset with **380k** sentences from **25** banks to evaluate LLM/PLM's ability to extract economics insights from textual data. Co-authored paper is accepted at NeurIPS 2025.
- Constructed an evaluation pipeline to assess retrievers in RAG for QA about credit agreements. Applied item extraction as data pre-processing to achieve **90.64%** retrieval accuracy.

• **Undergrad Research Assistant** — *Jun 2022 - April 2023*
Texera and Dr. Mandt's Lab, University of California, Irvine — Irvine, CA, USA
- Implemented elucidated diffusion model to generate spatial convective structures from Atmospheric information, achieving **100x** acceleration than traditional cloud-resolving models. Authored research poster is presented at the **Scientific ML Symposium 2023** in San Diego.
- Developed features (filtering, snapshotting, etc.) to enhance Texera's cloud-based workflow execution capabilities.
- Designed a version tracking algorithm, capturing code modifications to bolster system reproducibility.

SKILLS

• **Programming Languages:** Python, C/C++, JavaScript, TypeScript, SQL, Scala, HTML
• **Frameworks & Libraries:** PyTorch, Ray, Verl, LlamaIndex, Matplotlib, Pandas, Numpy, Flask, Juypter, Git, LaTex
• **Language:** Mandarin (Naive), English (Bilingual)
• **Passions:** Snowboarding, Basketball, Calligraphy, Photography