Xinjie Shen

frinkleko@gmail.com

Info

Georgia Institute of Technology (GaTech), Atalanta, USA — Aug. 2025 – Present

South China University of Technology (SCUT), Guangzhou, China — Sep. 2021 – 2025

Incoming PhD student in Machine Learning (First generation) **Homepage**: https://xinjie-shen.com **GPA:**3.77/4.0
Research Interests: Network&Connection, Data Mining, HCI, LLM Agents **Co-founder** of LLMQuant

Experience

Sparse Language Model for Searching — Nov. 2024 – Feb. 2025

Working with Zhichao Geng and Dr. Yang Yang at Amazon, Shanghai. Exploring the possibility of using sparse language model in searching. Released series of models recived **1.8M** downloads per month.

- (**SIGIR 2025 First author**) **Exploring ℓ_0 Sparsification for Inference-free Sparse Retrievers**
  Propose a new ℓ_0 sparsification loss to train the inference-free sparse retriever, which reduce the model size and achieve better performance under same FLOPS.

Investigating Overreliance issues on LLM — Sep. 2024 – Present

Working with Prof. Anthony Chen from UCLA and Prof. Sherry Wu from CMU. Investigating the over-reliance issues on LLM by designed large-scale user study. **One paper submitted(denoted with *).**

- *(**UIST 2025 Second author**) **Detecting Overreliance on Conversational LLM from Interaction Behaviors**
  Investigating the over-reliance issues on LLM by designing three users experiments including answering, summary and planning. Quantization and statics analysis used to investigate the correlations between user behaviors and over-reliance issues.

Explore Heterophily, Oversmoothing and Relations Modeling — Dec. 2023 – Oct.2024

Co-advised by Prof. Yujun Yan, Dartmouth College and Prof. Dawei Zhou, Virginia Tech. Explore possibility in modeling interaction to address heterophily and oversmoothing issues and improve upper bound on networks. **One paper submitted(denoted with *).**

- (**TMLR 2025 First author**) **Exploring the Potential of Complex-Valued GNN in Modeling Hidden Edge Relations**
  First propose to integrate complex-valued into GNN and asymmetric message passing by hidden edge relations modeling. Fresh insights and new direction on GNN design. Optimization on interaction leads to superior performance on networks.

Automatous R&D LLM Agent and Quantitative Finance Strategy in MSRA — Jan. 2024 – Aug.2024

Working with Dr. Jiang Bian, Microsoft Research Asia. Imagine a world where the R&D process is fully automated, and the LLM can automatically generate analysis results of every proposed ideas and propose new ideas. Project we led is open-sourced in: Github/RD-Agent**(5k+ Stars)**. Two preprint about it are released. Check our demo here. Highest honor **"Star of Tomorrow"** awarded by MSRA.

- (**Preprint Co-First author**) **Towards Data-Centric Automatic R&D**
  Defining the concept of Data-Centric Automatic R&D, providing both benchmarks and evaluation metrics for the LLM agent.
- (**Preprint Core author**) **Collaborative Evolving Strategy for Automatic Data-Centric Development**
  Propose a novel collaborative evolving strategy for automatic data-centric research and development with knowledge graph memory.

Multi-relational Geometric Information Interaction Research Group — Apr. 2023 – Dec. 2023

Advised by Prof. Danyang Wu, Xi'an Jiaotong University. Research on the graph neural network, geometric representation learning and their applications for RNA, protein, and etc. **Two paper accepted, two papers are under review (denoted with *).**

- (**TNNLS 2024 Co-First author**) **Cross-view Approximation on Grassmann Manifold for Multiview Clustering**
  Propose a novel framework for multi-view clustering that leverages the Grassmann manifold to approximate an orthogonal indicator matrix from multiple graph and feature views. Introduced adaptive weights to handle view inconsistency to optimize the objective.
- (**WWW 2024 Co-First author**) **FinReport: Explainable Stock Earnings Forecasting via News Factor Analyzing Model**
  Propose an explainable stock earning framework via news factor analyzing model. Stock earning foresting and explainable earning module are built via aggregating and utilizing news factor and numeric stock factor. A detailed report will be generated by LLM.
- *(**TKDE 2024 First author**) **NP^{2}L: Negative Pseudo Partial Labels Extraction for Graph Neural Networks**
  Propose an unsupervised framework to exact negative relationship between nodes, based on pseudo partial labels, and augment the original graph into a signed graph. SOTA performance on both link prediction and node classification tasks.
- (**WWW 2024 Second author**) **Simple Multigraph Convolution Networks**
  Propose a simple but effective multigraph convolution networks, focusing on exacting edge-level and subgraph-level credible fusion.

CTR/CVR Prediction& Smart Bidding Systems Developing In Tec-Do — Sep. 2022 – Jan. 2024

Data Mining and Machine Learning Engineer in Tec-Do (Future Unicorn), BIC Business Intelligence Center. **Awarded as BEST Project of the year. Two patents granted, two patents published and two patents under review.**

- Build CTR (Click-through rate)/CVR (Convert-through rate) predication model though GBDT and GNN for the Tec-Do Ads DSP platform, defeating 3rd party (Cusper) SaaS services. Achieved ROI 1.4x from 0.7x.
- Use modified DQN-like model to implement smart bidding module under constraint budget. 30% ROI improvement.
- Apply negative pseudo partial labels to sampling negative samples for recall model on bipartite graph.
- Build workflow generating real-person model picture for e-commerce based on diffusion model.

♡ Selected Awards

National Scholarship Highest scholarship awarded by the Chinese governments (**1%**) 2022, 2024
Star of Tomorrow Highest honor awarded by Microsoft Research Asia (**< 1%**) 2024
Project of the Year Awarded by Tec-Do for the DSP Platform 2023
TaiHu Innovation Prize Highest scholarship awarded by the WuXi city governments (**1%**) 2022
1st Prize Baidu Paddle Paddle Cup (**1%**) 2021
2nd Prize Intel Cup Undergraduate Electronic Design Contest (**8%**) 2022
1st Prize First price scholarship awarded by SCUT (**1%**) 2022

Leadership

I am also the co-founder of LLMQuant, an open-sourced community, which aims to explore the potential of LLM in the quantitative finance filed. Our product includes the QuantPedia and chatbot reshaping the learning and inquiring process.
I am also the chairman and the founder of the Artificial Intelligence Association (AIA) in SCUT. We have more than 200 members and 20+ projects, multiple interest-oriented groups, and occasional talks/salons.

Skills

- Programming Languages: Python, C++, SQL, JavaScript (TypeScript)
- Maths: Calculus, Linear Algebra, Complex Functions, Probability, Statistics, Convex Optimization, Graph Theory, Group Theory
- Platform&Tools: Windows, Linux; VSCode, Git, Docker, Nginx, LaTeX, Markdown
- Frameworks: Pytorch (Geometric), DGL, Stable Baselines3, Gym, Pandas (Polars), GBDTs (Catboost etc.), Visualization (Matplotlib, Seaborn, Plotly etc.), Scikit-learn, Numpy, Scipy, OpenCV, Flask, Pybind, Cython

Patents

1. Advertisement click prediction method based on random gradient attack, **CN116757748B**
2. Multi-task combined prediction method for few-sample advertisement, **CN117035873B**
3. Neural network system based on multidimensional data exploration, **CN117314531A**
4. Electronic commerce platform system based on artificial intelligence, **CN117196788A**
5. A method for patterned clothing fusion based on detail reconstruction, *(Under Review)*
6. An advertisement bidding system based on interval exploration, *(Under Review)*
7. An Ad Delay Conversion Prediction Method Based on Dual Window and Tree Modeling, *(Under Review)*

Miscellaneous

- Hobby: Anime, Literature (Novels and Poem), Badminton, Running
- In my spare time, I have keen interests in visualization communication design and VI system.
- I am extremely ethusiastic about teaching with vivid visualization and frequent language interaction.