Haitian (Hilton) Jiang

📞 +1 (332)-323-2919 ✉ haitian.jiang@nyu.edu 🔗 linkedin.com/in/htjiang 👤 haitian-jiang.github.io

🎓 Education

New York University   **Sep. 2023 – Present**
Ph.D. in Computer Science   *New York, NY*
Advisor: Prof. Jinyang Li & Prof. Aurojit Panda
Research Interest: Machine Learning, Systems for Machine Learning, Graph Neural Networks

Fudan University   **Sep. 2019 – Jun. 2023**
B.S. in Data Science (Summa Cum Laude)   *Shanghai, China*
Advisor: Prof. Zengfeng Huang

University of California, San Diego (UCSD)   **Sep. 2021 – Dec. 2021**
Exchange Student, Computer Science and Engineering   *La Jolla, CA*

📖 Publication

- **MuseGNN: Interpretable and Convergent Graph Neural Network Layers at Scale.**
  Haitian Jiang, *Renjie Liu, Xiao Yan, Zengfeng Huang, Zhenkun Cai, Minjie Wang, David Wipf.*
  Under review for ICLR 2025.

- **DiskGNN: Bridging I/O Efficiency and Model Accuracy for Out-of-Core GNN Training.**
  Renjie Liu, Yichuan Wang, Xiao Yan, Zhenkun Cai, Minjie Wang, ***Haitian Jiang***, *Bo Tang, Jinyang Li.*   **SIGMOD 2025.**

- **FreshGNN: Reducing Memory Access via Stable Historical Embeddings for Graph Neural Network Training.**
  Kezhao Huang*, ***Haitian Jiang*** (Equal Contribution), *Minjie Wang, Guangxuan Xiao, David Wipf, Xiang Song, Quan Gan, Zengfeng Huang, Jidong Zhai, Zheng Zhang.*   **VLDB 2024.**

- **Simplifying and Empowering Transformers for Large-Graph Representations.**
  Qitian Wu, Wentao Zhao, Chenxiao Yang, Hengrui Zhang, Fan Nie, ***Haitian Jiang***, *Yatao Bian, Junchi Yan.*   **NeurIPS 2023.**

💻 Research & Intern Experiences

Amazon Web Services   **May. 2024 – Present**
Applied Scientist Intern | Machine Learning Systems   *Santa Clara*
Advisor: Yida Wang & Zhen Jia
  - **Semantic Equivalence of Single Device and Distributed Training**

Systems Group, New York University   **Sep. 2023 – Present**
Research Assistant | Machine Learning Systems   *New York*
Advisor: Prof. Jinyang Li & Prof. Aurojit Panda
  - **Automatic Searching for Kernel Fusion and Scheduling**

DGL Group, Amazon Web Services (AWS) AI Lab   **Jun. 2022 – Aug. 2023**
Applied Scientist Intern | Machine Learning Theory and System   *Shanghai*
Advisor: David Wipf & Minjie Wang
  - **FreshGNN**: Graph Neural Network (GNN) training system with historical cache
    - Accelerated and scaled up GNN training on large graphs with a selective historical cache to reduce computation and data movement.
    - Designed managing policies based on gradients and staleness for cached nodes.
  - **MuseGNN**: Interpretable and Convergent Graph Neural Network Layers at Scale
    - Expanded a widely-used unfolded GNN framework to incorporate offline sampling into the architecture-inducing energy function design.
    - Demonstrated the model possess convergence properties of the bi-level optimization process.
    - Empirically extended the unfolded GNN framework to large graphs and achieved SOTA performance for homogeneous graph models applied to the largest open graph dataset.

TRAIL Lab, Fudan University   Jan. 2022 – Oct. 2022

Research Assistant | Graph Neural Network   *Shanghai*

Advisor: Prof. Zengfeng Huang

- **Scalable Graph Neural Networks using Subgraph Summarization**
  - Proposed a sub-graph sampling method that can cooperate with graph coarsening algorithms to reduce the size of sub-graphs by 30%.
  - Scaled up GNN training: implemented both pre-processing and training on multi-GPU; reduced the training time on papers100M from over 10 minutes to 27s for one epoch.

McAuley Lab, University of California, San Diego   Oct. 2021 – Aug. 2022

Research Assistant | Recommender System, Natural Language Processing   *La Jolla*

Advisor: Prof. Julian McAuley

- **Recommendation System with Faithful Textual Explanation**
  - Evaluated current text-based recommenders through aspect-based sentiment analysis and found the text cannot faithfully explain ratings as popular and positive aspects dominate the text.
  - Designed a metric of faithfulness for recommendation explanation by learning and comparing the mapping from explanations to predicted ratings in order to measure the issue.
  - Designed and implemented a new pipeline that first generate text explanations of recommendations and then use the text to predict faithful ratings.

Keen Lab, Tencent Technology   Feb. 2021 – Jul. 2021

Software Engineer | Reverse Engineering   *Shanghai, China*

- Collected, reproduced and exploited vulnerabilities (CVE) of IoT devices; conducted reverse-engineering on the firmware of devices.
- Implemented a crawler system using Scrapy, PostgreSQL, and message queue (RMQ) to automatically gather information of the products and download firmware.

🖳 Teaching Experience

TA: Honors Analysis of Algorithms, CSCI-GA.3110, Fall 2024, New York University

TA: Advanced Computer Graphics, CSCI-GA.2274, Spring 2024, New York University

TA: Big Data and Machine Learning Systems, CSCI-GA.3033(077), Spring 2024, New York University

TA: Advanced Big Data Analytics, DATA130014, Spring 2023, Fudan University

TA: Algorithm and Data Structures (Honor), DATA130023H, Fall 2022, Fudan University

🏛 Academic Services

Reviewer: NeurIPS 2024, ICLR 2025, AISTATS 2025

Organizer: NYU MLSys Seminar

🏅 Honors & Awards

Henry M. MacCracken Fellowship	**2023**
China National Scholarship (top 1%)	**2021**
National Second Prize (top 3%), CUMCM	**2020**
(Contemporary Undergraduate Mathematical Contest in Modeling)	
Second Prize Scholarship of Fudan University (top 10 %)	**2020**

🔨 Skills

Languages: Mandarin (Native speaker), English (Proficient)
Programming: Python, C/C++, CUDA, Triton, Go, MATLAB, R, SQL, PHP, Bash
Frameworks: Linux, Git, PyTorch, GPU programming, OpenMP, Spark, Docker, Web Crawler