Yilun Jin | Curriculum Vitae

Huawei Lianqiu Lake Research & Development Center – Shanghai – China

+86-18101885710, +852-65837148
yilun.jin@connect.ust.hk, jyl.jal123@gmail.com • kl4805.github.io

Education

- **Ph.D. , Dept. of CSE, HKUST** — **Hong Kong SAR**
  Computer Science and Engineering — *9.2019–5.2025*
  Advisor: Prof. Qiang Yang, Prof. Kai Chen.
- **B. S. , School of EECS, Peking University** — **Beijing, China**
  Computer Science, GPA: 3.7/4.0 — *9.2015–7.2019*
  Advisor: Prof. Guojie Song
  Selected courses: Advanced Algebra, Mathematical Analysis, Convex Optimization, Compiler Design.
- **B. Economics, National School of Development, Peking University** — **Beijing, China**
  Economics, GPA: 3.7/4.0 — *9.2016–7.2019*
  Selected courses: Intermediate Macro- & Micro-Economics, Econometrics, Financial Economics, Money and Banking, Special Topics on Reform in China
- **University of California, San Diego** — **La Jolla, USA**
  Exchange Student, GPA: 3.93/4.0 — *9.2017–12.2017*
  Courses: Computer Operating System, Machine Learning, Recommender System and Web Mining

Experiences

- **Huawei, 2012 Labs** — **Shanghai, China**
  Research & Development Engineer (Gifted Young Award) — *6.2025–*
  I work under the topic of AI-assisted Research & Development, and more specifically, various aspects on LLM for coding and agentic coding, such as building agentic applications, continue pre-training & post-training of LLMs, and synthetic training & evaluation data.
- **Amazon Stores Foundational AI** — **Palo Alto, USA**
  Applied Science Intern, advised by Dr. Zheng Li and Dr. Chenwei Zhang — *9.2023–4.2024*
  Research on a benchmark to evaluate LLMs on a wide-range of e-commerce-related applications. With this benchmark, we host the Amazon KDD Cup Competition 2024: Multi-task Online Shopping Challenge for LLMs.
- **Peking University, China** — **Beijing, China**
  Research Assistant, advised by Prof. Guojie Song — *9.2018–9.2019*
  Research on a wide range of topics through graph representation learning, including temporal aware, community aware and memory adaptive network embeddings. Several topics on Graph Neural Networks are also involved. See *Publications* for details.

Publications

(* stands for *Equal Contribution*.)

Conference Papers

1. **Shopping MMLU: A Massive Multi-Task Online Shopping Benchmark for Large Language Models** — 12.2024
   Yilun Jin, *Zheng Li, Chenwei Zhang, Tianyu Cao, Yifan Gao, Pratik Sridatt Jayarao, Mao Li, Xin Liu, Ritesh Sarkhel, Xianfeng Tang, Haodong Wang, Zhengyang Wang, Wenju Xu, Jingfeng Yang, Qingyu Yin, Xian Li, Priyanka Nigam, Yi Xu, Kai Chen, Qiang Yang, Meng Jiang, Bing Yin*
   In the 38th Conference on Neural Information Processing Systems (NeurIPS), Datasets and Benchmarks Track, 2024.

2. **Transferable Graph Structure Learning for Graph-based Traffic Forecasting across Cities** — 8.2023
   Yilun Jin, *Kai Chen, Qiang Yang*
   In the 29th SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2023.

3. **Selective Cross-city Transfer Learning for Traffic Prediction via Source City Region Re-weighting** — 8.2022
   Yilun Jin, *Kai Chen, Qiang Yang*
   In the 28th SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2022.

4. **Theoretically Improving Graph Neural Networks via Anonymous Walk Graph Kernels** — 4.2021
   *Qingqing Long**, ***Yilun Jin****, *Yi Wu**, *Guojie Song*
   In the Web Conference (TheWebConf, a.k.a. WWW), 2021.

5. **Graph Structural-topic Neural Network** — 8.2020
   *Qingqing Long**, ***Yilun Jin****, *Guojie Song, Yi Li, Wei Lin*
   In the 26th SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2020.

6. **EPNE: Evolutionary Pattern Preserving Network Embedding** — 6.2020
   *Junshan Wang**, ***Yilun Jin****, *Guojie Song, Xiaojun Ma*
   In the 24th European Conference on Artificial Intelligence (ECAI), 2020.

7. **GraLSP: Graph Neural Networks with Local Structural Patterns** — 2.2020
   Yilun Jin, *Guojie Song, Chuan Shi*
   In the 34th AAAI Conference on Artificial Intelligence, 2020.

8. **Unlocking Full Efficiency of Token Filtering in Large Language Model Training** — 4.2026
   Di Chai, Pengbo Li, Feiyuan Zhang, ***Yilun Jin****, Han Tian, Kaiqiang Xu, Binhang Yuan, Dian Shen, Junxue Zhang, Kai Chen*
   To appear in the 14th International Conference on Learning Representations (ICLR), 2026

9. **MixNet: A Runtime Reconfigurable Optical-Electrical Fabric for Distributed Mixture-of-Experts Training** — 9.2025
   Xudong Liao, Yijun Sun, Han Tian, Xinchen Wan, ***Yilun Jin*** *, Zilong Wang, Zhenghang Ren, Xinyang Huang, Wenxue Li, Kin Fai Tse, Zhizhen Zhong, Guyue Liu, Ying Zhang, Xiaofeng Ye, Yiming Zhang, Kai Chen*
   In the ACM SIGCOMM 2025 Conference, 2025

10. **Exploiting Student Parallelism for Low-latency GPU Inference of BERT-like Models in Online Services** — 8.2025
    Weiyan Wang, ***Yilun Jin****, Yiming Zhang, Victor Junqiu Wei, Han Tian, Li Chen, Jinbao Xue, Yangyu Tao, Di Wang, Kai Chen*

In the 31st SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2025

11. **Achieving Fairness Generalizability for Learning-based Congestion Control with Jury** *3.2025*
*Han Tian, Xudong Liao, Decang Sun, Chaoliang Zeng, **Yilun Jin**, Junxue Zhang, Xinchen Wan, Zilong Wang, Yong Wang, Kai Chen*
In the European Conference on Computer Systems (EuroSys), 2025

12. **Efficient Decentralized Federated Singular Vector Decomposition** *8.2024*
*Di Chai, Junxue Zhang, Liu Yang, **Yilun Jin**, Leye Wang, Kai Chen, Qiang Yang*
In the 2024 USENIX Annual Technical Conference (ATC), 2024

13. **Understanding Communication Characteristics of Distributed Training** *6.2024*
*Wenxue Li, Xiangzhou Liu, Yuxuan Li, **Yilun Jin**, Han Tian, Zhizhen Zhong, Guyue Liu, Ying Zhang, Kai Chen*
In the 8th Asia-Pacific Workshop on Networking (APNet), 2024

14. **Accelerating Privacy-Preserving Machine Learning With GeniBatch** *5.2024*
*Xinyang Huang, Junxue Zhang, Xiaodian Cheng, Hong Zhang, **Yilun Jin**, Shuihai Hu, Han Tian, Kai Chen*
In the European Conference on Computer Systems (EuroSys), 2024.

15. **MDP: Model Decomposition and Parallelization of Vision Transformer for Distributed Edge Inference** *12.2023*
*Weiyan Wang, Yiming Zhang, **Yilun Jin**, Han Tian, Li Chen*
IEEE International Conference on Mobility, Sensing, and Networking (MSN), 2023.

16. **Scalable and Efficient Full-Graph GNN Training for Large Graphs** *6.2023*
*Xinchen Wan, Kaiqiang Xu, Xudong Liao, **Yilun Jin**, Kai Chen, Xin Jin*
In the ACM Conference on Management of Data (SIGMOD), 2023.

17. **GraphMSE: Efficient Meta-path Selection in Semantically Aligned Feature Space for Graph Neural Networks** *2.2021*
*Yi Li, **Yilun Jin**, Guojie Song, Chuan Shi, Zihao Zhu, Yiming Wang*
In the 35th AAAI Conference on Artificial Intelligence, 2021.

18. **Domain Adaptive Classification on Heterogeneous Information Networks** *7.2020*
*Shuwen Yang, Guojie Song, **Yilun Jin**, Lun Du*
In the 29th International Joint Conference on Artificial Intelligence (IJCAI-PRICAI), 2020.

19. **Active Domain Transfer on Network Embedding** *4.2020*
*Lichen Jin, Yizhou Zhang, Guojie Song, **Yilun Jin***
In the Web Conference (TheWebConf, a.k.a. WWW), 2020.

20. **Hierarchical Community Structure Preserving Network Embedding: A Subspace Approach** *11.2019*
*Qingqing Long, Yiming Wang, Lun Du, Guojie Song, **Yilun Jin**, Wei Lin*
In the 28th ACM International Conference on Information and Knowledge Management (CIKM), 2019.
Best Research Paper Runner-up

21. **DANE: Domain Adaptive Network Embedding** *8.2019*
*Yizhou Zhang, Guojie Song, Lun Du, Shuwen Yang, **Yilun Jin***
In the 28th International Joint Conference on Artificial Intelligence (IJCAI). 2019.

Journal Papers

22. **Federated Learning without Full Labels: A Survey** *5.2023*
***Yilun Jin**, Yang Liu, Kai Chen, Qiang Yang*
In IEEE Data Engineering Bulletin, 2023

23. **SecureBoost: A Lossless Federated Learning Framework**
*Kewei Cheng, Tao Fan, **Yilun Jin**, Yang Liu, Tianjian Chen, Qiang Yang* *5.2021*
In IEEE Intelligent Systems, 2021

24. **Deep Convolutional Neural Network based Medical Concept Normalization**
*Guojie Song, Qingqing Long, Yi Luo, Yiming Wang, **Yilun Jin*** *9.2020*
In IEEE Transactions on Big Data, 2020

Technical and Personal skills

- **Mathematics Background:** Discrete Mathematics, Advanced Algebra, Convex Optimization, Mathematical Analysis, Probability and Statistics.
- **Programming Languages:** C, C++, Python (TensorFlow, PyTorch), Java, LaTeX, SQL
- **Language Skills:** Chinese (native speaker), English (proficient with speaking 26 in TOEFL, analytical writing 5.0 in GRE), Japanese (N2), Cantonese (beginner)
  - I take charge of writing in most of the publications listed above.
- **General Skills:** Self motivation. Work well in a team.

Teaching

- **Teaching Assistant**, *Practice of Programming in C++*, Peking University, Spring 2019
- **Teaching Assistant**, *COMP4631: Computer and Communication Security*, The Hong Kong University of Science and Technology, Spring 2020
- **Teaching Assistant**, *COMP5631: Cryptography and Security*, The Hong Kong University of Science and Technology, Fall 2020
- **Teaching Assistant**, *COMP6211G: Federated Learning*, The Hong Kong University of Science and Technology, Spring 2021
- **Teaching Assistant**, *COMP3511: Operating Systems*, The Hong Kong University of Science and Technology, Fall 2021, Fall 2022

Services

- **Area Chair**: KDD 2026.
- **Main Organizer**: Amazon KDD Cup Challenge 2024.
- **Vice Local Chair**: APNet 2023.
- **Conference Program Committee Member/Reviewer**:
  - **2020**: TheWebConf (Poster & Demo), ICONIP
  - **2021**: AAAI, IJCAI, NeurIPS, NeurIPS Datasets and Benchmarks Track, ICONIP
  - **2022**: ICLR, IJCAI, ICML, TheWebConf (Poster & Demo), NeurIPS, NeurIPS Datasets and Benchmarks Track, LoG
  - **2023**: AAAI, ICLR, IJCAI, TheWebConf (Industry), KDD, ICML, ECML-PKDD, NeurIPS, NeurIPS Datasets and Benchmarks Track, LoG
  - **2024**: AAAI, ICLR, SDM, TheWebConf, IJCAI, ICML, KDD, ECML-PKDD, NeurIPS, NeurIPS Datasets and Benchmarks Track
  - **2025**: AAAI, ICLR, KDD, ICML, NeurIPS, NeurIPS Datasets and Benchmarks Track, LoG
  - **2026**: AAAI, ICLR, KDD
- **Journal Reviewer**: IEEE BigData, IEEE T-ITS, Neural Networks, TMLR, Information Fusion, IEEE IoTJ, IEEE TKDE.
- **Secondary Reviewer**:

- **2020**: AAAI, ACL, NeurIPS
- **2023**: SIGSPATIAL

Awards

- **Excellent Reviewer** (Top 20%), *ACM SIGKDD Conference*, 2025.
- **Student Travel Award**, *ACM SIGKDD Conference*, 2023.
- **2021 Best Paper Award**, *IEEE Intelligent Systems*, 2021
- **Top Reviewer**, *NeurIPS*, 2022
- **Outstanding Reviewer** (Top 10%), *ICML*, 2022
- **AAAI Student Scholarship**, *AAAI*, 2020
- **Best Research Paper Runner-up**, *CIKM Research Track*, 2019
- **Award for Research Excellence**, *Peking University*, *2018*
- **Merit Student**, *Peking University*, *2017*
- **Huawei Scholarship**, *Peking University*, *2017*
- **Yu Minhong Scholarship for overseas exchange**, *Peking University*, *2017*
- **Founder Scholarship**, *Peking University*, *2016*
- **Award for Academic Excellence**, *Peking University*, *2016*