Pengfei Hu

213-548-1411 phu9@stevens.edu humphreyhuu.github.io/
Address: 1 Castle Point Terrace, GS-433, Hoboken, NJ 07030

Education

Stevens Institute of Technology **Sep. 2023 – Present**
Ph.D. in Computer Science (GPA: 3.95/4) *Hoboken, NJ*

- **Ph.D. Advisor:** Dr. Yue Ning;
- **Research Interests:** Clinical AI, Domain Generalization, Large Language Models, Graph Neural Networks;
- Teaching Assistant for CS-584: Natural Language Processing on Fall 2024 and Spring 2025.

University of Southern California **Sep. 2021 – May 2023**
M.S. in Industrial System Engineering (GPA: 4/4) *Los Angeles, CA*

- Research Assistant advised by Dr. Sze-chuan Suen working on monitoring of diabetes via wearable sensor devices.

Nanjing University of Finance and Economics **Sep. 2016 – May 2020**
B.A. in Economic Statistics (Rank: 1 / 145, GPA: 3.63/4) *Nanjing, China*

- Working on relative deprivation in the digital era (one Q1 publication) supervised by Dr. Qinghai Li.

Work Experience

Oak Ridge National Laboratory **May 2025 – Aug. 2025**
Graduate Research Intern, Mentor: Ming Fan *Oak Ridge, TN*

- Utilized the Frontier supercomputer server to deploy parallel training on over 10 compute nodes via DDP and mpi4py;
- **Project:** Provide multi-reservoir inflow predictive support for water management in the Upper Colorado River Basin;
- **Paper:** Two conference workshop papers in ICDM 2025 and AAAI 2026 and extended journal submissions.

Alibaba Group, Inc. **Mar. 2021 – Jul. 2021**
Data Analyst Intern *Hangzhou, China*

- Tracked 20GB delivery data via Hive SQL to evaluate performance like freight volume and penetration of shipment to filter potential counties, calculated ROI and compared confident intervals to allocate ads budget efficiently by 10%;
- Summarized high-risk users and matched them with behaviour-based features (fake Mac-Address, pageview status, online duration etc.), constructed and tuned best model LightGBM by Python for prediction with 87% precision rate and saved 7% budget for fraud detection of risk control comparing with $100k+ cost of last quarter.

Other Internships

- Product Analyst Intern at Tencent (Apr. 2020 – Aug. 2020, *Shenzhen, China*);
- Risk Data Analyst Intern at Didi Global (Dec. 2019 – Apr. 2020, *Chengdu, China*);
- Actuarial Analyst Intern at PwC Consulting (Jan. 2019 – Mar. 2019, *Shenzhen, China*).

Publications

(The google scholar link can be found here, * stands for equal contribution.)

1. **Pengfei Hu**, Xiaoxue Han, Fei Wang, Yue Ning, "UdonCare: Hierarchy Pruning for Unseen Domain Discovery in Predictive Healthcare", Under Review, 2025.
2. **Pengfei Hu**, Chang Lu, Fei Wang, Yue Ning, "Bridging Stepwise Lab-Informed Pretraining and Knowledge-Guided Learning for Diagnostic Reasoning", Under Review by Journal, 2025.
3. **Pengfei Hu**, Fan Ming, Xiaoxue Han, Chang Lu, Yue Ning, Dan Lu "Hydrological Domain-Conditioned Modulation for Cross-Reservoir Inflow Prediction", *Accepted by AAAI Workshop on AI for Environmental Science*, 2026. **Oral.**
4. Xiaoxue Han, **Pengfei Hu**, Jun-En Ding, Chang Lu, Feng Liu, Yue Ning, "Interpretable and Interactable Predictive Healthcare with Knowledge-Enhanced Agentic Causal Discovery", *In Findings of EMNLP 2025*, 2025.
5. **Pengfei Hu**, Fan Ming, Xiaoxue Han, Chang Lu, Wei Zhang, Hyun Kang, Yue Ning, Dan Lu "Adaptive Graph Learning with Transformer for Multi-Reservoir Inflow Prediction", *IEEE International Conference on Data Mining (ICDM) 2025 Workshop on Data Mining in Earth System Science*, 2025.
6. Eric Yang*, **Pengfei Hu***, Xiaoxue Han, Yue Ning, "MPLite: Multi-Aspect Pretraining for Mining Clinical Health Records", *2024 IEEE International Conference on Big Data (IEEE BigData)*, 2024.
7. Hsin-Ling Hsu, Cong-Tinh Dao, ..., **Pengfei Hu**, ..., Chenwei Wu (8^{th}/15 authors in total), "MEDPLAN: A Two-Stage RAG-Based System for Personalized Medical Plan Generation", *In Proceedings of ACL 2025 Industry Track*, 2025.

Technical Skills

Programming Languages: Python, SQL, Java, R, Latex.
Deep Learning Framework: Pytorch, PyG, Keras.