

WEIJIAN LI

✉ weijianli2021@u.northwestern.edu | 📞 (+1) 7738072950 | 🏠 1100 Church Street, Evanston, IL, USA 60201
🔍 Google Scholar Profile | 🔗 LinkedIn Profile

EDUCATION

Northwestern University, United States

Sept 2019 – Present

Candidate for Doctor of Philosophy, Computer Science

GPA: 3.85/4.00 Advisor: Orrington Lunt Professor, Han Liu

Master of Science, Computer Science

GPA: 4.00/4.00

University of Edinburgh, United Kingdom

Sept 2016 – Jul 2019

Bachelor of Science, Artificial Intelligence & Computer Science

GPA: 3.99/4.00 (1st Honors)

RESEARCH INTERESTS

I work broadly on large-scale deep learning and artificial intelligence. My recent focus is on improving the efficacy of large foundation models (pre-training LLMs, Vision Transformers) and applying these techniques for scientific discoveries (DNABERT-2, astrophysical time series). My research interests include:

- Artificial General Intelligence (LLMs and multi-modality model training)
- Foundation Models for Scientific Discoveries (DNA and astrophysical time series)

WORK EXPERIENCE

-
- **Applied Scientist Intern, Amazon** June 2024 – Sept 2024
Designed and developed an LLM-based time series model to reduce customer risks.
 - **Applied Scientist Intern, Amazon** June 2023 – Sept 2023
Designed and developed a machine learning pipeline using an LLM for a classification problem.
 - **Research Intern, Boerboel Trading** Aug 2021 – Nov 2021
Designed and developed high-frequency trading infrastructure and strategies.
 - **Teaching Assistant, Data Science Capstone, Northwestern University** Dec 2020 – Mar 2021
Developed data science projects and delivered lectures on time series prediction pipelines.
 - **Research Assistant, Northwestern University** Aug 2019 – Dec 2019
Developed *PredicT*, a time series prediction platform using XGBoost and deployed it on a Kubernetes cluster.
 - **Web Developer, University of Edinburgh Library Team** June 2017 – Sept 2017
Developed a web game to crowdsource human descriptions of the library’s digital collections.

PUBLICATIONS

Artificial General Intelligence

- [1] Tim Tsz-Kit Lau*, **Weijian Li***, Chenwei Xu, Han Liu, and Mladen Kolar. “**Adaptive Batch Size Schedules for Distributed Training of Language Models with Data and Model Parallelism**”. In: *Conference on Parsimony and Learning (CPAL)* (2025). URL: <https://arxiv.org/abs/2412.21124>.
- [2] Tim Tsz-Kit Lau, **Weijian Li**, Chenwei Xu, Han Liu, and Mladen Kolar. “**Communication-Efficient Adaptive Batch Size Strategies for Distributed Local Gradient Methods**”. In: *arXiv preprint* (2024). submitted to AISTATS 2025. URL: <https://arxiv.org/abs/2406.13936>.

- [3] Jerry Yao-Chieh Hu, Pei-Hsuan Chang, Haozheng Luo, Hong-Yu Chen, **Weijian Li**, Wei-Po Wang, and Han Liu. “**Outlier-Efficient Hopfield Layers for Large Transformer-Based Models**”. In: *Proceedings of the 41st International Conference on Machine Learning*. Ed. by Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp. Vol. 235. Proceedings of Machine Learning Research. PMLR, 21–27 Jul 2024, pp. 19123–19152. URL: <https://proceedings.mlr.press/v235/hu24a.html>.

Foundation Models for Scientific Discoveries

- [4] Zhihan Zhou, Yanrong Ji, **Weijian Li**, Pratik Dutta, Ramana V Davuluri, and Han Liu. “**DNABERT-2: Efficient Foundation Model and Benchmark For Multi-Species Genomes**”. In: *The Twelfth International Conference on Learning Representations*. 2024. URL: <https://openreview.net/forum?id=oMLQB4EZE1>.
- [5] Zhi Zhang, **Weijian Li**, and Han Liu. “**Multivariate Time Series Forecasting By Graph Attention Networks With Theoretical Guarantees**”. In: *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*. Ed. by Sanjoy Dasgupta, Stephan Mandt, and Yingzhen Li. Vol. 238. Proceedings of Machine Learning Research. PMLR, Feb. 2024, pp. 2845–2853. URL: <https://proceedings.mlr.press/v238/zhang24g.html>.
- [6] Dennis Wu*, Jerry Yao-Chieh Hu*, **Weijian Li***, Bo-Yu Chen, and Han Liu. “**STanHop: Sparse Tandem Hopfield Model for Memory-Enhanced Time Series Prediction**”. In: *The Twelfth International Conference on Learning Representations*. 2024. URL: <https://openreview.net/forum?id=6iwg437CZs>.
- [7] Chenwei Xu, Yu-Chao Huang, Jerry Yao-Chieh Hu, **Weijian Li**, Ammar Gilani, Hsi-Sheng Goan, and Han Liu. “**BiSHop: Bi-Directional Cellular Learning for Tabular Data with Generalized Sparse Modern Hopfield Model**”. In: *Proceedings of the 41st International Conference on Machine Learning*. Ed. by Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp. Vol. 235. Proceedings of Machine Learning Research. PMLR, 21–27 Jul 2024, pp. 55048–55075. URL: <https://proceedings.mlr.press/v235/xu241.html>.
- [8] Alex Reneau, Jerry Yao-Chieh Hu, Chenwei Xu, **Weijian Li**, Ammar Gilani, and Han Liu. “**Feature Programming for Multivariate Time Series Prediction**”. In: *International Conference on Machine Learning* (2023). arXiv: 2306.06252 [cs.LG]. URL: <https://arxiv.org/abs/2306.06252>.
- [9] Tong Xie, Yuwei Wan, **Weijian Li**, Qingyuan Linghu, Shaozhou Wang, Yalun Cai, Han Liu, Chunyu Kit, Clara Grazian, and Bram Hoex. “**Interdisciplinary Discovery of Nanomaterials Based on Convolutional Neural Networks**”. In: *NeurIPS 2022 AI for Science: Progress and Promises*. 2022. URL: <https://openreview.net/forum?id=jrx4vTYqRzJ>.
- [10] Lingzhi Wang, Xiangmin Shen, **Weijian Li**, Zhenyuan Li, R Sekar, Han Liu, and Yan Chen. “**Incorporating gradients to rules: Towards lightweight, adaptive provenance-based intrusion detection**”. In: *arXiv preprint arXiv:2404.14720* (2024). Accepted by Network and Distributed System Security Symposium 2025.

Papers Under Review

- [2] Tim Tsz-Kit Lau, **Weijian Li**, Chenwei Xu, Han Liu, and Mladen Kolar. “**Communication-Efficient Adaptive Batch Size Strategies for Distributed Local Gradient Methods**”. In: *arXiv preprint* (2024). submitted to AISTATS 2025. URL: <https://arxiv.org/abs/2406.13936>.
- [11] **Weijian Li** and Han Liu. ***RDAS: A Low Latency and High Throughput Raw Data Engine for Machine Learning Systems***. submitted to ICML 2025.
- [12] **Weijian Li**, Haozheng Luo, Chenwei Xu, and Han Liu. ***SWGA: A Distributed Hyperparameter Search Method for Time Series Prediction Models***. submitted to ICML 2025.
- [13] **Weijian Li**, Stephen S Cheng, Lining Mao, Jigyasa Kumari, Alex Pyo, Mehak Kawatra, Jialong Li, Jiayi Wang, Ammar Gilani, Jingya Xun, Jui-Hui Chung, Jerry Yao-Chieh Hu, and Han Liu. ***A Benchmark Study For Limit Order Book (LOB) Models and Time Series Forecasting Models on LOB Data***. submitted to ICML 2025.

*Equal contribution

SKILLS

Pandas, Numpy, PyTorch, Hugging Face, PyTorch Lightning, Scikit-learn, Flask, React, Docker, Kubernetes, AWS, Hadoop, Python, Java, C++