# Chang Liu

changl8@andrew.cmu.edu | changliu-drpatient.github.io

## Education

***School of Computer Science***

*Carnegie Mellon University, Ph.D. Student in Machine Learning* 08/2023 - 05/2028(est.)

- Advisor: Prof. Artur Dubrawski.
- Research Interests: LLM post-training, Agentic AI, machine learning in healthcare.
- Skills: LLM fine-tuning; Prompt engineering; Harness engineering.
- Courses: Advanced Introduction to Machine Learning, Intermediate Statistics, Advanced Machine Learning: Theory and Methods, Advanced Deep Learning, Machine Learning for Large Datasets.

***Yao Class, Institute for Interdisciplinary Information Sciences (IIIS)*** Established by Prof. Andrew C. Yao

*Tsinghua University, B.Eng. in Computer Science* 08/2019 - 06/2023

- GPA: **3.91/4.00**.
- TOEFL: **120/120**. Reading: 30, Listening: 30, Speaking: 30, Writing: 30.
- GRE: **340/340**. Quant: 170, Verbal: 170, Writing: 5.
- Mathematics Courses: Calculus, Linear Algebra, Abstract Algebra, Mathematics for Computer Science, Mathematics for Artificial Intelligence, Probability and Statistics.
- Computer Science Courses: Machine Learning, Reinforcement Learning, Computational Biology, Computer Vision, Deep Learning, Natural Language Processing, Introduction to Databases, Data Mining, Quantum Computer Science, Introduction to Robotics, Algorithm Design, Theory of Computation.

## Industry Experience

***TikTok: Video Recommendation***

*Research Intern* 05/2026 - 08/2026

- Model user preference on PGC (professionally generated content) to boost recommendation performance.

## LLM Research Experience

***Experiential Abstractions in LLMs*** 05/2026 - Now

*LLM post-training. Reinforcement learning. verl. Retrieval-augmented LLMs.* Carnegie Mellon University

- Develop a novel framework for improving LLM problem-solving abilities by first extracting abstractions from training, and then absorbing them via in-context learning or RL. (Build on ByteDance's verl library for RL.)
- Improve mathematical reasoning abilities compared to naive GRPO, even without a strong teacher for abstraction extraction.
- Submitted to EMNLP. Paper link here.

***Lightweight LLM agents for MLE*** 05/2025 - 08/2025

*Coding agents. AWS strands. Prompt engineering. Harness engineering* Carnegie Mellon University

- Developed a lightweight LLM agent for machine learning engineering (MLE).
- Adopt the CodeAct paradigm (draft and revise) and invoke mini-swe-agent as a tool.
- Use AWS's strands library for building the agents.
- Project link here.

***Depth pruning and emergent abilities of pruned LLMs*** 12/2024 - 09/2025

*LLM pruning. LLM inference. Huggingface. Benchmarking* Carnegie Mellon University

- Proposed a novel depth-based LLM pruning methodology.
- Systematically benchmarked the emergent abilities of aggressively pruned LLMs.
- Paper link here.

## Healthcare Research Experience

***Multimodal outbreak detection in hospitals*** 10/2023 - Now

*Multimodal ML. ML for healthcare.* Carnegie Mellon University

- Develop a multimodal machine learning framework for hospital outbreak detection that integrates signals from MALDI-TOF mass spectra, antimicrobial resistant patterns, and electronic health records (EHR).
- Enable cost-effective outbreak detection that can reduce dependency on whole genome sequencing gold standard.

***Multimodal disease target identification through a probabilistic knowledge graph*** 09/2021 - 05/2023

*Multimodal ML. Graph Neural Networks. Computational biology.* Tsinghua University

- Develop a novel method of augmenting biological networks with literature evidence to construct a probabilistic knowledge graph.
- Develop a graph neural network to predict target candidates from the knowledge graph, achieving superior performance to state-of-the-art models in terms of accuracy (esp. on sparse data) and literature support for top novel predictions.
- Conduct bioinformatics analyses and cooperate with experimental validation of the identified colorectal cancer and melanoma targets.

## Publications

1. **Chang Liu**; Xinyu Li; Artur Dubrawski*.“Notes to Self: Can LLMs Benefit from Experiential Abstractions?” submitted to *EMNLP*, May 2026. paper link here.
2. **Chang Liu**; Jieshi Chen; Alexander J. Sundermann; Kathleen Shutt; Marissa P. Griffith; Lora Lee Pless; Lee H. Harrison; Artur Dubrawski*.“Towards Practical Multimodal Outbreak Detection,” *AMIA Annual Symposium*, May 2026.
3. **Chang Liu**; Jieshi Chen; Alexander J. Sundermann; Kathleen Shutt; Marissa P. Griffith; Lora Lee Pless; Lee H. Harrison; Artur Dubrawski*.“Exploring the Utility of MALDI-TOF Mass Spectrometry and Antimicrobial Resistance in Hospital Outbreak Detection,” *AMIA Amplify Informatics Conference*, February 2026.
4. **Chang Liu**†; Arjun Choudhry†; Yifu Cai; Nina Żukowska; Mononito Goswami; Artur Dubrawski. “Depth as a Scaling Vector: Simple Pruning and Evaluation of Emergent Abilities in Pruned LLMs,” *NeurIPS 2025 Workshop on Evaluating the Evolving LLM Lifecycle: Benchmarks, Emergent Abilities, and Scaling*, September 2025.
5. Arjun Choudhry†; **Chang Liu**†; Nina Żukowska; Yifu Cai; Mononito Goswami; Artur Dubrawski. “LayerMerge: Modality-Agnostic Depth Pruning for Efficient Foundation Model Deployment,” *NeurIPS 2025 Workshop on Efficient Reasoning*, September 2025.
6. Mikołaj Piórczyński, Wojciech Łapacz, Xinyu Li, **Chang Liu**, Abby Turner, Artur Dubrawski. “From Aggregation to Guidance: Strategies for Personalized Federated Fine-Tuning of Foundation Models” *NeurIPS 2025 Workshop on Unifying Representations in Neural Models*, September 2025.
7. **Chang Liu**; Jieshi Chen; Lee H. Harrison; Artur Dubrawski*. “Bridging the Utility Gap Between MALDI-TOF and WGS for Affordable Outbreak Cluster Detection,” *The AHLI Conference on Health, Inference, and Learning (CHIL)*, April 2025.
8. **Chang Liu**; Jieshi Chen; Lee H. Harrison; Artur Dubrawski*. “Multimodal Structure Preservation Learning,”*arXiv preprint*, October 2024.
9. **Chang Liu**†; Kaimin Xiao†; Cuinan Yu†; Yipin Lei†;...; Dan Zhao*; Fengfeng Zhou*; Haidong Tang*; Jianyang Zeng*. “A Probabilistic Knowledge Graph Approach for Target Identification,” *PLOS Computational Biology*, April 2024.

10. **Chang Liu**†; Cuinan Yu†; Yipin Lei†;...; Dan Zhao*; Fengfeng Zhou*; Jianyang Zeng*. "Improving Target-disease Association Prediction through a Graph Neural Network with Credibility Information," proceedings of the *Pacific Symposium on Biocomputing*, January 2023.

## Honors & Awards

- **Comprehensive Merit Award (7/32)**, *Tsinghua University* 2022
- **Comprehensive Merit Award (6/32)**, *Tsinghua University* 2021
- **Excellence Award for Volunteering Services**, *Tsinghua University* 2020
- **Freshmen Scholarship**, *Tsinghua University* 2019
- **University Full Scholarship for Future Scholars**, *Tsinghua University* 2019

## Academic Service

- **Peer Review:** Reviewed 16 academic papers. NeurIPS 2026 (1), NeurIPS UniReps Workshop 2025 (2), AMIA Amplify Informatics Conference 2026 (5), AMIA Annual Symposium (8).