

Yingtao Luo

Tel: +1 4129547651 • Pittsburgh, PA • Email Address: yingtaol@andrew.cmu.edu • [Linkedin](#) • [Google Scholar](#)

SUMMARY

- Ph.D. candidate in Machine Learning with M.S. degrees in Machine Learning and Computer Science, specializing in Large Language Models, Agentic AI, Reinforcement Learning, and Biomedical AI for Health Applications.
- Experienced in large-scale ML model development and LLM fine-tuning (PEFT, SFT, CoT, RLHF, RLAIF, RAG).
- Skilled in handling multimodal biomedical data such as electronic health records (EHRs), imaging, and multi-omics.

EDUCATION

Carnegie Mellon University, PA, USA

Ph.D. in Machine Learning & Public Policy 08/2022 – 02/2027 (exp.)

M.S. in Machine Learning Research (halfway Ph.D. in Machine Learning) 08/2022 – 05/2025

University of Washington, WA, USA

M.S. in Computer Science and Systems 10/2020 – 06/2022

Huazhong University of Science and Technology, China

B.S. in Applied Physics 09/2015 – 06/2019

EXPERIENCES

Research Fellow, Center of Machine Learning for Health, CMU (Pittsburgh, PA) 01/2024 – now

- Developed a deep reinforcement learning agent for real-time decision-making of heart-transplant donor acceptance.
- Improved simulated patient survival by +2.7 years and +8% one-year survival, achieving 96% physician agreement.
- Built an agentic LLM interface for interpretable recommendations, with 86% overall clinician satisfaction.
- Recognized by NIH grant extension (5 years) and CMLH Translational Fellowship (2 years).

PhD Student Researcher, Machine Learning Department, CMU (Pittsburgh, PA) 09/2023 – now

➤ *Project 1: Medical Deep Research Co-Pilot (Agentic Multimodal Reasoning System)*

- Built a web-based multimodal [Medical Deep Research](#) system, integrating EHRs, imaging, notes, and literature.
- Developed RAG-based evidence pipelines, reducing hallucinations by 97% and improving factual grounding.
- Delivered 5/5 LLM evaluation and 4.8/5 clinician satisfaction, validating real-world usability.

➤ *Project 2: Medical Reasoning LLM for Clinical Decision Support*

- Fine-tuned Qwen3-8B models with GPT4o-generated CoT, RLHF and RLAIF to enhance clinical reasoning.
- Created a data engine integrating synthetic and real EHRs for scalable clinical LLM corpus construction.
- Outperformed models under 70B in diagnostic and treatment-plan optimization benchmarks like HealthBench.

➤ *Project 3: A Foundation Model for Capturing Long-Range Gene Context*

- Built scLong, a billion-parameter transformer pretrained on 48M single-cell samples and ontology graphs.
- Outperformed state-of-the-art models in drug response and gene perturbation prediction accuracy.

Research Intern, Decision Intelligence Group, Damo Academy, Alibaba (Seattle, WA) 2023.05 – 2023.08

- Built a Bayesian online multivariate imputation framework for streaming data, outperforming SOTAs.
- Designed physics-informed models for weather forecasting with infinite granularity and real-time scalability.

Research Intern, Machine Learning Group, Microsoft Research Asia (Beijing, China) 2021.09 – 2022.03

- Developed an interpretable neural differential operator for state-of-the-art time series forecasting.
- Contributed to the development of core Transformers and RL models in Qlib (30k★ GitHub).

PROFICIENT SKILLS

- **Programming:** Python (Pytorch, TensorFlow, DeepSpeed, LangChain, LlamaIndex, transformers, etc.), SQL, Spark
- **Machine Learning:** Large Language Models, Deep Learning, Reinforcement Learning, NLP, Multimodal Fusion, etc.
- **Tools:** Git, Linux, Docker, Kubernetes, Jupyter Notebook, VS Code, PyCharm, Weights & Biases, Azure, GCP Cloud
- **Biomedical Data:** Multimodal EHR (MIMIC-IV, ADNI, UK Biobank, Synthea), QA Corpora (PubMedQA, MedQA, BioASQ), Multi-Omics (PDB, Human Cell Atlas, AlphaFold), Biomedical Knowledge Graphs, HealthBench, etc.

AWARDS

- CMU Presidential Fellowship (2022), CMLH Translational Fellowship in Digital Health (2024)
- Finalist for Best Paper Award in AMIA Student Paper Competition (2025)

PAPERS

➤ *AI for Healthcare*

- Shao, R., Seraj, M. S.*, Zhao, K.*, Luo, Y.* (Equal Contribution), Li, L., Shen, B., Bates, A., Zhao, Y., Pan, C., Hightow-Weidman, L., Chakraborty, S., Dong., Y. (2025). LLM-Empowered Medical Patient Communication: A Data-Centric Survey From a Clinical Perspective. Accepted to AACL 2025.
- Li, Y., Zhu, H., Liu, C., Pan, M., Luo, Y.# (Corresponding Author). (2025). Demo: Orchestrating Large Language Model Agents and Resources for Medical Investigation. Accepted to NeurIPS2025 GenAI4Health Workshop.
- Luo, Y., Li, Y., Zhao, K., Shau, E., Au, I. (To be finalized). Towards Real-World Clinical Decision-Making Support: A state-of-the-art medical reasoning model. Experiments mostly done and in preparation for NEJM AI.
- Luo, Y., Padman, R., Skandari, R., Kilic, A. (2025). Prediction of Mortality After Adult Heart Transplantation Using the Updated UNOS Registry. Under review at JAMA Cardiology.
- Luo, Y., Skandari, R., Kilic, A., Padman, R. (2025). Simulated Patient Trajectories to Inform Decision-Making in Heart Transplantation. Under review at JAMA.
- Luo, Y., Skandari, R., Martinez, C., Kilic, A., Padman, R. (2025). Benchmarking Waitlist Mortality Prediction Through Time-to-Event Modeling using New UNOS Dataset. Proceedings of AMIA Annual Symposium. **Finalist for the Martin Epstein Award (Best Student Paper)**.
- Bai, D., Mo, S., Zhang, R., Luo, Y., Gao, J., Yang, J. P., Wu, Q., Singh, D., Rahmani, H., Amariuta, T., Grotjahn, D., Zhong, S., Lewis, N., Wang, W., Ideker, T., Xie, P., & Xing, E. (2024). scLong: A Billion-Parameter Foundation Model for Capturing Long-Range Gene Context in Single-Cell Transcriptomics. Submitted to Nat. Commun.
- Mahbub, S., Ellington, C., Alinejad, S., Wen, K., Luo, Y., Lengerich, B., & Xing, E. (2024). From One to Zero: RAG-IM Adapts Language Models for Interpretable Zero-Shot Predictions on Clinical Tabular Data. NeurIPS Adaptive Foundation Model Workshop.
- Deusel, J.*., Ellington, C.*., Luo, Y., Lengerich, B., Friederich, Xing, E. (2024). Contextualized Policy Recovery: Modeling and Interpreting Medical Decisions with Adaptive Imitation Learning. International Conference on Machine Learning (ICML).
- Luo, Y., Li, Z., Liu, Q., Zhu, J. (2024). Fairness without Demographics on Electronic Health Records. AAAI Spring Symposium on Clinical Foundation Models. **Contributed Talk (Top 20%)**.
- Luo, Y., Liu, Z., Liu, Q. (2022). Deep Stable Representation Learning on Electronic Health Records. IEEE International Conference on Data Mining (ICDM).

➤ *General AI Methodology*

- Luo, Y., Li, Z., Liu, Q., Zhu, J. (2025). Fairness without Demographics through Learning Graph of Gradients. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD).
- Luo, Y., Li, Z., Liu, Q., Zhu, J. (2025). Physics-Guided Learning of Meteorological Dynamics for Weather Downscaling and Forecasting. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD).
- Jiang, J.*., Zhang, P.*., Luo, Y.* (Equal Contribution), Li, C., Kim, J. B., Zhang, K., Wang, S., Kim, S., Yu, P. S. (2025). Improving Sequential Recommendations via Bidirectional Temporal Data Augmentation with Pre-Training. IEEE Transactions on Knowledge and Data Engineering (TKDE).
- Wu, Y., Luo, Y., Kong, X., Papalexakis, E. E., & Steeg, G. V. (2024). Your Diffusion Model is Secretly a Noise Classifier and Benefits from Contrastive Training. Advances in Neural Information Processing Systems (NeurIPS).
- Fang, S., Wen, Q., Luo, Y., Zhe, S., Sun, L. (2024). BayOTIDE: Bayesian Online Multivariate Time series Imputation with functional decomposition. International Conference on Machine Learning (ICML). **Spotlight Paper (Top 2%)**.
- Li, Z., Wang, L., Sun, X., Luo, Y., Zhu, Y., Chen, D., Luo, Y., Zhou, X., Liu, Q., Wang, L., Yu, J. (2023). GSLB: The Graph Structure Learning Benchmark. Advances in Neural Information Processing Systems (NeurIPS).
- Luo, Y.* (Equal Contribution), Liu, Q.*., Chen, Y., Hu, W., Tian, T., Zhu, J. (2023). Physics-Guided Discovery of Highly Nonlinear Parametric Partial Differential Equations. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) & NeurIPS AI for Science Workshop.

- Jiang, J.* , Zhang, P.* , Luo, Y., Li, C., Kim, J., Zhang, K., Wang, S., Xie, X., Kim, S. (2023). AdaMCT: Adaptive Mixture of CNN-Transformer for Sequential Recommendation. ACM International Conference on Information and Knowledge Management (CIKM).
- Liu, Q., Luo, Y., Wu, S., Zhang, Z., Yue, X., Jin, H., Wang, L. (2022). Reject-aware Multi-Task Network for Financial Credit Scoring. IEEE Transactions on Knowledge and Data Engineering (TKDE).
- Luo, Y., Xu, C., Liu, Y., Liu, W., Zheng, S., Bian, J. (2022). Learning Differential Operators for Interpretable Time Series Modeling. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD).
- Chen, Y., Luo, Y., Liu, Q., Xu, H., Zhang, D. (2022). Symbolic genetic algorithm (SGA) for discovering open-form partial differential equations. Physics Review Research.
- Zhang, Z.* , Luo, Y.* (Equal Contribution), Peng, H., Chen, Y., Liao, R.-Z., Zhao, Q. (2020). Deep spatial learning of polyamide nanofiltration membranes. Journal of Membrane Science & NeurIPS ML4Mol Workshop.
- Luo, Y.-T., Li, P.-Q., Li, D.-T., Peng, Y.-G., Geng, Z.-G., Xie, S.-H., Li, Y., Alù A. Zhu J., Zhu, X.-F. (2020). Probability-Density-Based Deep Learning Paradigm for the Design of Functional Metastructures. Research (Science Partner Journal) & NeurIPS ML4PS Workshop.
- Luo, Y., Liu, Q., Liu, Z. (2021). STAN: Spatio-Temporal Attention Network for the Next Location Recommendation. Proceedings of the Web Conference (WWW2021).