
RAN XU

E-mail: ran.xu@emory.edu | Webpage: ritaranx.github.io

Address: 400 Dowman Dr, Atlanta, GA 30322, USA

EDUCATION

Laney Graduate School, Emory University

Atlanta, GA, USA

Ph.D. in Computational Science and Informatics

Aug. 2021 - Present

- Ph.D. Advisors: [Dr. Carl Yang](#) and [Dr. Joyce Ho](#);
- **Thesis Topic:** Advancing Precision Medicine with Heterogeneous Healthcare Data
- **Research Interest:** Health Informatics, Natural Language Processing, Large Language Models, Graph Neural Networks

Department of Computer Science and Mathematics, Emory University

Atlanta, GA

B.S. in Computer Science; B.S. in Applied Mathematics

Aug. 2017 - May 2021

- GPA: 3.969/4.00;
- Research Assistant advised by [Dr. Jinho Choi](#) working on early detection of Alzheimer's Disease from patients' speech recordings.

INDUSTRY EXPERIENCE

Amazon.com, Inc.

Palo Alto, CA, USA

Applied Scientist Intern, Mentor: [Hui Liu](#)

May. 2024 - Oct. 2024

- **Project:** LLM Self-training for retrieval-augmented generation
- **Paper:** One arXiv paper [under review](#).

Meta Platforms, Inc. (formerly named Facebook, Inc.)

Menlo Park, CA, USA

Enterprise Engineer Intern, Mentor: [Zexi Zhang](#)

May. 2020 - Aug. 2020

- **Project:** Implemented a procurement platform from both front- and back-end for communication between employees and suppliers.

PUBLICATIONS

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

1. **Ran Xu**, Hui Liu, Sreyashi Nag, Zhenwei Dai, Yaochen Xie, Xianfeng Tang, Chen Luo, Yang Li, Joyce C Ho, Carl Yang, Qi He. "SimRAG: Self-Improving Retrieval-Augmented Generation for Adapting Large Language Models to Specialized Domains", Under Review, 2024.
2. Jiaming Shen, **Ran Xu**, Yennie Jun, Zhen Qin, Tianqi Liu, Carl Yang, Yi Liang, Simon Baumgartner, Michael Bendersky. "Boosting Reward Model with Preference-Conditional Multi-Aspect Synthetic Data Generation", Under Review, 2024.
3. **Ran Xu***, Wenqi Shi*, Yue Yu*, Yuchen Zhuang, Yanqiao Zhu, May D Wang, Joyce C Ho, Chao Zhang, Carl Yang. "BMRetriever: Tuning large language models as better biomedical text retrievers", In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2024.
4. Wenqi Shi*, **Ran Xu***, Yuchen Zhuang, Yue Yu, Jieyu Zhang, Hang Wu, Yuanda Zhu, Joyce Ho, Carl Yang, May D Wang. "EHRAgent: Code Empowers Large Language Models for Complex Tabular Reasoning on Electronic Health Records", In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2024.
5. Wenqi Shi*, **Ran Xu***, Yuchen Zhuang, Yue Yu, Hang Wu, Carl Yang, May D Wang. "MedAdapter: Efficient Test-Time Adaptation of Large Language Models towards Medical Reasoning", In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2024.
6. **Ran Xu**, Yiwen Lu, Chang Liu, Yong Chen, Yan Sun, Xiao Hu, Joyce C Ho, Carl Yang. "From Basic to Extra Features: Hypergraph Transformer Pretrain-then-Finetuning for Balanced Clinical Predictions on EHR", In *Proceedings of the Conference on Health, Inference, and Learning (CHIL)*, 2024.
7. **Ran Xu***, Wenqi Shi*, Yue Yu, Yuchen Zhuang, Bowen Jin, May D Wang, Joyce C Ho, Carl Yang. "RAM-EHR: Retrieval Augmentation Meets Clinical Predictions on Electronic Health Records", In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*, 2024. **Oral.**
8. **Ran Xu**, Hejie Cui, Yue Yu, Xuan Kan, Wenqi Shi, Yuchen Zhuang, Wei Jin, Joyce Ho, Carl Yang. "Knowledge-Infused Prompting Improves Clinical Text Generation with Large Language Models", In *Findings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL)*, 2024.

9. Ziyang Zhang, Hejie Cui, **Ran Xu**, Yuzhang Xie, Joyce C Ho, Carl Yang. “[TACCO: Task-guided Co-clustering of Clinical Concepts and Patient Visits for Disease Subtyping based on EHR Data](#)”, In *Proceedings of the 30th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2024.
10. **Ran Xu**, Yue Yu, Joyce C. Ho and Carl Yang. “[Weakly-supervised Scientific Document Classification via Retrieval-Augmented Multi-stage Training](#).” In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2023.
11. **Ran Xu**, Yue Yu, Hejie Cui, Xuan Kan, Yanqiao Zhu, Joyce C Ho, Chao Zhang, and Carl Yang. “[Neighborhood-regularized Self-Training for Learning with Few Labels](#).” In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI)*, 2023. **Oral**.
12. **Ran Xu**, Mohammed K Ali, Joyce Ho, Carl Yang. “[Hypergraph Transformers for EHR-based Clinical Predictions](#).”. In *Proceedings of the American Medical Informatics Association Informatics Summit (AMIA)*, 2023.
13. **Ran Xu**, Yue Yu, Chao Zhang, Mohammed K Ali, Joyce C Ho, and Carl Yang. “[Counterfactual and Factual Reasoning over Hypergraphs for Interpretable Clinical Predictions on EHR](#)”. In *Proceedings of the 2nd Machine Learning for Health symposium (ML4H)*, 2022. **Best Paper Award**.
14. Yue Yu, Rongzhi Zhang, **Ran Xu**, Jieyu Zhang, Jiaming Shen and Chao Zhang. “[Cold-Start Data Selection for Few-shot Language Model Fine-tuning: A Prompt-Based Uncertainty Propagation Approach](#).” In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.
15. Hejie Cui, Xinyu Fang, Zihan Zhang, **Ran Xu**, Xuan Kan, Xin Liu, Manling Li, Yangqiu Song, Carl Yang. “[Open Visual Knowledge Extraction via Relation-Oriented Multimodality Model Prompting](#)” In *Proceedings of the Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS)*, 2023.
16. Yue Yu, Xuan Kan, Hejie Cui, **Ran Xu**, Yujia Zheng, Xiangchen Song, Kun Zhang, Raziel Nabi, Ying Guo, Chao Zhang and Carl Yang. “[Deep DAG Learning of Effective Brain Connectivity for fMRI Analysis](#).” In *Proceedings of the 20th IEEE International Symposium on Biomedical Imaging (ISBI)*, 2023.
17. Xuan Kan, Zimu Li, Hejie Cui, Yue Yu, **Ran Xu**, Shaojun Yu, Zilong Zhang, Ying Guo, Carl Yang, “[R-Mixup: Riemannian Mixup for Biological Networks](#)”, In *Proceedings of the 29th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2023.
18. Hejie Cui, Jiaying Lu, Shiyu Wang, **Ran Xu**, ..., Carl Yang (12 authors in total), “[A Survey on Knowledge Graphs for Healthcare: Resources, Applications, and Promises](#)”, In *the 3rd Workshop on Interpretable Machine Learning in Healthcare*, 2023.
19. Hejie Cui, Xinyu Fang, **Ran Xu**, Xuan Kan, Joyce C Ho, Carl Yang. “[Multimodal fusion of EHR in structures and semantics: Integrating clinical records and notes with hypergraph and LLM](#)”, Under Review, 2024.

HONORS AND AWARDS

• Emory Global Diabetes Research Center Fellowship	2024
• Best Reviewer in KDD 2024	2024
• Best Paper Award at Machine Learning for Health 2022 (Top 2)	Nov. 2022
• Women in Natural Sciences Fellowship	Fall 2021
• The Chair’s Award for Achievements (Top 1)	2020, 2021
• Highest Honor in Honor thesis	Spring 2021
• Dean’s List	Fall 2017, Spring 2018, Fall 2018, Spring 2019

PROFESSIONAL SKILLS

-
- Programming language: Python, Java, MATLAB, Latex.
 - Deep learning frameworks: Pytorch, Keras.

SERVICES

-
- **Teaching Experience:**
 - Teaching Assistant for CS570: Data Mining. Spring 2023, Spring 2022
 - Teaching Assistant for CS253: Data Structures and Algorithms. Spring 2021, Fall 2021
 - **Conference and Journal Review:** ICML 2024, ICLR 2025, 2024; NeurIPS 2024, 2023; COLM 2024; MLHC 2024, 2023; CHIL 2024, 2023; LoG 2023; KDD 2024, 2023, 2022; WWW 2023, 2022; AMIA 2023; ACM TIST 2024; TBD 2022; TKDE 2021; ACL Rolling Reviews (bimonthly).