Jiaying Lu

Updated: 2025-04-18

Assistant Professor
Center for Data Science
Nell Hodgson Woodruff School of Nursing, Emory University
Suite 245, 1520 Clifton Rd, Atlanta, GA

jiaying.lu AT emory.edu
(+1) (404) 307-8738
lujiaying.github.io
Google Scholar

EDUCATION BACKGROUNDS

Ph.D. Computer Science & Informatics, Emory University, USA, 2019–2024
Dissertation: Learning Structured Knowledge from Real-World Data without Excessive Annotations
Advisor: Dr. Carl Yang

M.Eng. Electronics & Communication Engineering, Beijing University of Posts and Telecommunications, China, 2014–2017
Exchange student at Aalto University, Finland, Sep 2015–Jan 2016

B.Sc. Information Engineering, Beijing University of Posts and Telecommunications, China, 2010–2014

ACADEMIC APPOINTMENTS

2024– Center for Data Science, Nell Hodgson Woodruff School of Nursing, Emory University.
Research-Track Assistant Professor

HONORS & AWARDS

2024 **Young Scientist Excellence Award** (4th-place winner). *MidSouth Computational Biology and Bioinformatics Society*. Atlanta, GA, USA.

2022 NSF Registration Award. *31st ACM International Conference on Information and Knowledge Management*. Atlanta, GA, USA.

2022 NSF Student Travel Award. *28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. Washington, D.C., USA.

2022 Registration Award. *Georgia Statistics Day 2022*. Athens, GA, USA.

2021–24 Emory Professional Development Support Funds. *The James T. Laney School of Graduate Studies*. Atlanta, GA, USA.

GRANTS

Ongoing Research Support

2024–26 *Incorporating Structured Knowledge in Large Language Models for Precision Healthcare* (Lu)
Startup Grant
Emory Nursing Data Science Foundation
Role: Principal Investigator

2025 *Understanding Generative AI's Impact to Users from Various Religious Background* (Lu)
Seed Grant
Atlanta Interdisciplinary Artificial Intelligence Network
Role: Principal Investigator

PUBLICATIONS

Equal contribution denoted with *, corresponding authors with ✉.

Journal Articles

(J3) **Jiaying Lu**, Ran Xiao, Xiao Hu, and Duc H. Do. Artificial intelligence in cardiac telemetry. *BMJ Heart*, 2025b. doi: 10.1136/heartjnl-2024-323947

(J2) **Jiaying Lu**, Xiangjue Dong, and Carl Yang. Weakly supervised concept map generation through task-guided graph translation. *IEEE Transactions on Knowledge and Data Engineering*, 35(10): 10871–10883, March 2023. doi: 10.1109/TKDE.2023.3252588

(J1) Wenjing Ma, **Jiaying Lu**, and Hao Wu. Cellcano: supervised cell type identification for single cell atac-seq data. *Nature Communications*, 14(1):1864, April 2023. ISSN 2041-1723. doi: 10.1038/s41467-023-37439-3 (2023 ASA SSGG **Distinguished Student Paper Award**)

Conference Papers

(C14) **Jiaying Lu***✉, Stephanie R. Brown*, Songyuan Liu, Shifan Zhao, Kejun Dong, Del Bold, Michael Fundora, Alaa Aljiffry, Alex Fedorov, Jocelyn Grunwell, and Xiao Hu. Early risk prediction of pediatric cardiac arrest from electronic health records via multimodal fused transformer. In *47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (EMBC), 2025

(C13) Saurabh Kataria, Ran Xiao, Timothy Ruchti, Matthew T Clark, **Jiaying Lu**, Randall Lee, Jocelyn Grunwell, and Xiao Hu. Continuous cardiac arrest prediction in icu using ppg foundation model. In *47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (EMBC), 2025

(C12) Mingchen Shao, Youjeong Kang, Xiao Hu, Hyunjung Gloria Kwak, Carl Yang, and **Jiaying Lu**✉. Mining social determinants of health for heart failure patient 30-day readmission via large language models. In *The 20th World Congress on Medical and Health Informatics - Poster*, (MedInfo), 2025 (Student Supported by Emory Undergraduate Research Conference Presentation Grant)

(C11) Yuhao Xu, **Jiaying Lu**, Sirui Ding, Xiao Hu, and Carl Yang. Are foundation models useful for electrocardiogram analysis? a multi-task benchmark with comprehensive evaluations and insightful findings. In *The 20th World Congress on Medical and Health Informatics*, (MedInfo), 2025

(C10) Guanchen Wu, Linzhi Zheng, Han Xie, Zhen Xiang, **Jiaying Lu**, Darren Liu, Delgersuren Bold, Bo Li, Xiao Hu, and Carl Yang. Large language model empowered privacy-protected framework for phi annotation in clinical notes. In *The 20th World Congress on Medical and Health Informatics*, (MedInfo), 2025

(C9) Bo Xiong, **Jiaying Lu**, Yuqicheng Zhu, and Carl Yang. Mash: Maximal separating poincaré hyperplanes for hierarchical imbalanced learning. In *The ACM Web Conference 2025 - Short Paper*, (WWW), 2025

(C8) Yuzhang Xie, **Jiaying Lu**, Joyce Ho, Fadi Nahab, Xiao Hu, and Carl Yang. Promptlink: Leveraging large language models for cross-source biomedical concept linking. In *47th International ACM SIGIR Conference on Research and Development in Information Retrieval - Short Paper*, (SIGIR), July 2024b

(C7) **Jiaying Lu**, Shifan Zhao, Wenjing Ma, Hui Shao, Xiao Hu, Yuanzhe Xi, and Carl Yang. Uncertainty-aware pre-trained foundation models for patient risk prediction via gaussian process. In *The ACM Web Conference 2024 (HealthDay Track)*, (WWW), May 2024. doi: 10.1145/3589335.3651456 (A short version is selected for the **Young Scientist Excellence Award** at MCBIOS'24.)

(C6) **Jiaying Lu***, Yongchen Qian*, Shifan Zhao, Yuanzhe Xi, and Carl Yang. Mug: A multimodal classification benchmark on game data with tabular, textual, and visual fields. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, (Findings-EMNLP), December 2023. doi: 10.18653/v1/2023.findings-emnlp.354

(C5) **Jiaying Lu**, Jiaming Shen, Bo Xiong, Wengjing Ma, Staab Steffen, and Carl Yang. Hiprompt: Few-shot biomedical knowledge fusion via hierarchy-oriented prompting. In *46th International ACM SIGIR Conference on Research and Development in Information Retrieval - Short Paper*, (SIGIR), April 2023. doi: 10.1145/3539618.3591997

(C4) Xiangjue Dong, **Jiaying Lu**, Jianling Wang, and James Caverlee. Closed-book question generation via contrastive learning. In *17th Conference of the European Chapter of the Association for Computational Linguistics*, (EACL), May 2023. doi: 10.18653/v1/2023.eacl-main.230

(C3) **Jiaying Lu** and Carl Yang. Open-world taxonomy and knowledge graph co-learning. In *4th Conference on Automated Knowledge Base Construction*, (AKBC), November 2022

(C2) Hejie Cui, **Jiaying Lu**, Yao Ge, and Carl Yang. How can graph neural networks help document retrieval: A case study on cord19 with concept map generation. In *44th European Conference on Information Retrieval - Short Paper*, (ECIR), April 2022. doi: 10.1007/978-3-030-99739-7_9

(C1) **Jiaying Lu** and Jinho D. Choi. Evaluation of Unsupervised Entity and Event Salience Estimation. In *Proceedings of the 34rd International Florida Artificial Intelligence Research Society Conference*, (FLAIRS), May 2021. doi: 10.32473/flairs.v34i1.128482

Workshop Papers

(W5) Zhaoliang Chen, Cheng Ding, Nirbhay Modhe, **Jiaying Lu**, Carl Yang, and Xiao Hu. Adapting a generative pretrained transformer achieves sota performance in assessing diverse physiological functions using only photoplethysmography signals: A gpt-ppg approach. In *AAAI 2024 Spring Symposium on Clinical Foundation Models*, Feburary 2024

(W4) **Jiaying Lu***, Jinmeng Rao*, Kezhen Chen, Xiaoyuan Guo, Yawen Zhang, Baochen Sun, Carl Yang, and Jie Yang. Evaluation and enhancement of semantic grounding in large vision-language models. In *AAAI 2024 Workshop on Responsible Language Models*, (ReLM), February 2024 (**Spotlight**)

(W3) Zhexiong Liu, Jing Zhang, **Jiaying Lu**, Wenjing Ma, and Joyce Ho. Logicprpbank: A corpus for logical implication and equivalence. In *AAAI 2024 Workshop on AI for Education*, (AAAI-AI4ED), February 2024c

(W2) Hejie Cui, **Jiaying Lu**, Shiyu Wang, Ran Xu, Wenjing Ma, Shaojun Yu, Yue Yu, Xuan Kan, Chen Ling, Joyce Ho, et al. A survey on knowledge graphs for healthcare: Resources, applications, and promises. In *ICML 2023 Workshop on Interpretable Machine Learning in Healthcare*, (IMLH@ICML), July 2023

(W1) **Jiaying Lu**, Xin Ye, Yi Ren, and Yezhou Yang. Good, better, best: Textual distractors generation for multi-choice VQA via policy gradient. In *CVPR 2022 Workshop on Open-Domain Retrieval Under a Multi-Modal Setting*, (O-DRUM@CVPR), June 2022. doi: 10.1109/CVPRW56347.2022.00539

Preprint Reports

(P10) Shuyue Jiang, Wenjing Ma, Runze Yan, Chang Su, and **Jiaying lu**✉. Chronological age prediction based on dna methylation and clinical features, 2025

(P9) Runze Yan*, Guanlin Dai*, Yufen Lin, Yuhua Wu, **Jiaying Lu**, Xiao Hu, and Canhua Xiao. Predict epigenetic aging at an early stage: A case study to predict post-radiotherapy aging in head and neck cancer patients, 2025

(P8) **Jiaying Lu**, John R. Bowblis, Shuang Li, Yong-Fang Kuo, Jennifer Heston-Mullins, James S. Goodwin, and Huiwen Xu. Patterns and predictors of allowing visitors in nursing homes during the covid-19 pandemic, 2025a

(P7) Yuzhang Xie, Xu Han, Xiao Hu Ran Xu, **Jiaying Lu**, and Carl Yang. Hypkg: Hypergraph-based knowledge graph contextualization for precision healthcare, 2024a

(P6) Simon Liu, Ziyang Zhang, Runze Yan, Wei Wu, Carl Yang, and **Jiaying Lu**✉. Measuring spiritual values and bias of large language models. arXiv:2410.11647, 2024b

(P5) Shifan Zhao, **Lu, Jiaying**, Ji Yang, Edmond Chow, and Yuanzhe Xi. Efficient two-stage gaussian process regression via automatic kernel search and subsampling. arXiv:2405.13785, 2024

(P4) Runze Yan, Hanqi Luo, **Jiaying Lu**, Darren Liu, Terryl J Hartman, and Xiao Hu. Dietai24: A trustworthy dietary assessment framework with customizable multimodal large language models, 2024

(P3) Darren Liu, Cheng Ding, Delgersuren Bold, Monique Bouvier, **Jiaying Lu**, Benjamin Shickel, Craig S. Jabaley, Wenhui Zhang, Soojin Park, Michael J. Young, Mark S. Wainwright, Gilles Clermont, Parisa Rashidi, Eric S. Rosenthal, Laurie Dimisko, Ran Xiao, Joo Heung Yoon, Carl Yang, and Xiao Hu. Evaluation of general large language models in contextually assessing semantic concepts extracted from adult critical care electronic health record notes. *arXiv:2401.13588*, 2024a

(P2) Guangji Bai, Zheng Chai, Chen Ling, Shiyu Wang, **Jiaying Lu**, Nan Zhang, Tingwei Shi, Ziyang Yu, Mengdan Zhu, Yifei Zhang, Carl Yang, Yue Cheng, and Liang Zhao. Beyond efficiency: A systematic survey of resource-efficient large language models. *arXiv:2401.00625*, 2024

(P1) Chen Ling, Xujiang Zhao, **Jiaying Lu**, Chengyuan Deng, Can Zheng, Junxiang Wang, Tanmoy Chowdhury, Yun Li, Hejie Cui, Tianjiao Zhao, et al. Beyond one-model-fits-all: A survey of domain specialization for large language models. *arXiv:2305.18703*, 2023 (Cited by **2024 Economic Report of the President of USA**)

Patents

2023 Jinmeng Rao, **Jiaying Lu**, Kezhen Chen, Jie Yang, Yawen Zhang, Xiaoyuan Guo. "A Simulation-Feedback Framework to Mitigate Hallucinations in Multimodal Generative Models". *Filed 2023. X-52882-00-US*.

INVITED TALKS

2024 "Leveraging Language Models for Predicting Cardiac Arrest with Multimodal Electronic Health Records in Pediatric ICU". **Research Seminar.** Children's Healthcare of Atlanta. Atlanta, GA, USA. Aug 18, 2024.

2023 "Retrieval Augmented Generation for Biomedical Entity Alignment via Large Language Models". **Doctoral Symposium.** 2023 Conference on Health, Inference, and Learning. Boston, MA, USA, Jun 24, 2023.

2022 "Learning to Construct Structured Knowledge". **Data Science Forum.** School of Data Science, The Chinese University of Hong Kong. Palo Alto, CA, USA. Dec 17, 2022.

TEACHING EXPERIENCES

Courses Taught

2025 Co-Instructor. *NRSG 753 Advanced Computation for Nursing Research I*. Emory University.

2021 Teaching Assistant. *CS570 Data Mining*. Emory University.

2020 Teaching Assistant. *CS253 Data Structure and Algorithms*. Emory University.

2020 Teaching Assistant. *CS170 Introduction to Computer Science I*. Emory University.

Doctoral Student Advisement

2024– Andrew Lu, Ph.D. in Nursing AI, Emory University (*Thesis Committee*).

- Early Detection of Occlusion Myocardial Infarction through Responsible AI.

Undergraduate Advisement

2024–25 Simon Liu, B.S. in Computer Science, Emory University (*Honor Thesis Committee*).

- Representing Human Contact as Hypergraph: A Deep Learning Approach.

2024–25 Tung Dinh, B.S. in Computer Science, Emory University (*Honor Thesis Committee*).

2024–25 Mingchen Shao, B.S. in Computer Science, Emory University (*Research Advisor*).

- Large Language Model Enhanced Automated Phenotype Extraction from Clinical Notes.

PROFESSIONAL EXPERIENCES

2023 **AI Residency**. Mountain View, CA, USA. May–Sep 2023.
Mineral Research Team, **Google X**.
Supervisors: Dr. Jinmeng Rao, Dr. Kezhen Chen.

- Data-centric approach for mitigating hallucination in multimodal large language models.

2022 **Applied Scientist Intern**. Sunnyvale, CA, USA. May–Aug 2022.
Auto Machine Learning Team, **Amazon.com**.
Supervisors: Dr. Andrew Borthwick, Nick Erickson.

- Multi-modal ensemble models for large-scale product classification problems.

2018–19 **Visiting Scholar**. Phoenix, AZ, USA. Aug 2018–Jan 2019.
Active Perception Group, **Arizona State University**.
Advisor: Dr. Yezhou Yang.

- Research Interests: Visual Question Answering, Multimodal Learning.

2018 **NLP Engineer**. Beijing, China. Jan–Jul 2018.
Search Engine Group. **Xiaomi**.

- Word segmentation, named entity recognition, unsupervised synonym pairs mining.

2016–17 **Machine Learning Engineer Full-Time & Intern**. Beijing, China. Oct 2016–Dec 2017.
Search Advertising Group. **Baidu**.

- Ads content generation, Query-Ads matching.

2016 **Machine Learning Engineer Intern**. Beijing, China. Jul–Sep 2016.
Growth Hacking Team. **Meituan**.

- Gradient boosted regression trees for new customer coupon distribution prediction.

2016 **Machine Learning Engineer Intern**. Beijing, China. Mar–Jun 2016.
Online Food Delivery Group. **Baidu**.

- Logistic regression model for customer CTR prediction.

2014–15 **Software Engineer Intern**. Beijing, China. Sep 2014–Feb 2015.
Beijing YunJiang Technology Co., Ltd.

- Web & mobile application backend development.

PROFESSIONAL MEMBERSHIPS

2024– Associate Member, *Society of Critical Care Medicine*.

2024– Member, *Georgia Center for Diabetes Translation Research Data Repository*.

2023– Member, *Association for Computing Machinery*.

2024– Member, *Special Interest Group on Information Retrieval*.

SERVICES

Academic Event Organizations

2025 Poster Award Judge. *The 2025 Southeast Regional Clinical and Translational Science Conference*.

2022 Web Developer Student Lead. *28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (**KDD**).

Journal Peer Reviews

2023– The Journal of Agricultural, Biological and Environmental Statistics (**JABES**).

2022– The IEEE Transactions on Neural Networks and Learning Systems (**TNNLS**).

2021 The IEEE Transactions on Knowledge and Data Engineering (**TKDE**).

2021 The IEEE Transactions on Big Data (**TBD**).

Conference Peer Reviews

2023,25 The Conference on Machine Learning for Health (**ML4H**)

2022,25 The ACM Conference on Knowledge Discovery and Data Mining (**KDD**).

2024– The Association for Computational Linguistics Rolling Review (**ACL Rolling Review**).

2023 The ACM Special Interest Group on Information Retrieval (**SIGIR**).

2021-22 The Conference on Empirical Methods in Natural Language Processing (**EMNLP**).

2023 The IEEE International Conference on Data Mining (**ICDM**).

2023 The IEEE International Conference on Bioinformatics and Biomedicine **(BIBM)**.

2022- The International World Wide Web Conference (**WWW**).

2022 The AAAI Conference on Artificial Intelligence (**AAAI**).

2021 The ACM International Conference on Information and Knowledge Management (**CIKM**).

2021 The International Joint Conference on Natural Language Processing (**IJCNLP**).

2020 The International Joint Conferences on Artificial Intelligence (**IJCAI**).

2019-23 The IEEE International Conference on Robotics and Automation (**ICRA**).

Symposium & Workshop Peer Reviews

2024 AAAI Spring Symposium on Clinical Foundation Models.

2023 NeurIPS Workshop on Socially Responsible Language Modelling Research (**SoLaR**).

2023 ICLR Workshop on Trustworthy Machine Learning for Healthcare (**TML4H**).

2023– ACL SIGEDU Workshop on Innovative Use of NLP for Building Educational Applications (**BEA**).

2021– ICML Workshop on Interpretable Machine Learning in Healthcare (**IMLH**).

2021– ICCV Workshop on Computer Vision for Automated Medical Diagnosis (**CVAMD**).