Zirui (Ray) Liu

3400 Edloe St. Houston, Texas, 77027

☎979-985-7904 ✉zl105@rice.edu **G**Google scholar **in**LinkedIn profile ⌂Home page

RESEARCH INTEREST

– **Machine Learning Algorithm & System**: Large Language Models (LLMs) efficiency & scalability;
– **Applications**: Large-scale scientific and social computing problems.

EDUCATION

Rice University Houston, TX
- Ph.D. Candidate in Computer Science, Co-advised by
Dr. Xia (Ben) Hu and Prof. Vladimir Braverman *Aug. 2021 – May. 2024 (Expected)*

Texas A&M University (Transfer Out) College Station, TX
- *Ph.D. Student in Computer Science, Advised by Dr. Xia (Ben) Hu* *Aug. 2019 – Aug. 2021*

Harbin Institute of Technology Harbin, China
- *Electrical Engineering, Master of Engineering* *Sep. 2016 – Jun. 2018*

Harbin Institute of Technology Harbin, China
- *Electrical Engineering, Bachelor of Science* *Sep. 2012 – Jul. 2016*

RESEARCH & INDUSTRY EXPERIENCES

Research Intern at Linkedin Sunnyvale, CA
- *Advertisement AI team, advised by Mr. Aman Gupta.* *May. 2023 – Aug. 2023*
  ○ Improving the inference efficiency of language models via dynamic sparsity.

Research Intern at Meta Sunnyvale, CA
- *Capacity Engineering and Analysis team.* *May. 2022 – Aug. 2022*
  ○ Incorporating semi-supervised learning for improving advertisement models.

Research Intern at Samsung Research America Remote
- *Advertisement AI team, advised by Dr. Li Li and Dr. Rui Chen.* *May. 2021 – Aug. 2021*
  ○ Incorporating the graph structure into the App embeddings for improving the click-through-rate.

Research Intern at 4Paradigm Beijing, China
- *Advised by Mr. Xiawei Guo.* *Mar. 2019 – Aug. 2019*
  ○ Developed the automated machine learning system for tabular data.

PUBLICATION

1. [**ICLR'24 Submission**] **Zirui Liu**, Zhaozhuo Xu, Beidi Chen, Yuxin Tang, Jue Wang, Kaixiong Zhou, Xia Hu, Anshumali Shrivastava *"Compress, Then Prompt: Improving Accuracy-Efficiency Trade-off of LLM Inference with Transferable Prompt "*, ArXiv Preprint & Under review, PDF

2. [**NeurIPS'23**] **Zirui Liu**, Guanchu Wang, Shaochen Zhong, Zhaozhuo Xu, Daochen Zha, Ruixiang Tang, Zhimeng Jiang, Kaixiong Zhou, Vipin Chaudhary, Shuai Xu, and Xia Hu*"Winner-Take-All Column Row Sampling for Memory Efficient Adaptation of Language Model "*, The 37th Conference on Neural Information Processing Systems, 2023, PDF

3. [**KDD MLG'23**] **Zirui Liu**, Zhimeng Jiang, Shaochen Zhong, Kaixiong Zhou, Li Li, Rui Chen, Soo-Hyun Choi and Xia Hu. "*Editable Graph Neural Network for Node Classifications*", The 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining on Mining and Learning with Graphs Workshop, 2023. PDF

4. [**TMLR'23**] **Zirui Liu**, Kaixiong Zhou, Zhimeng Jiang, Li Li, Rui Chen, Soo-Hyun Choi, and Xia Hu. "*DSpar: Embarrassingly Simple Strategy for Efficient GNN training and inference via Degree-based Sparsification*", Transactions on Machine Learning Research, 2023. PDF

5. [**ICML'23**] **Zirui Liu**, Shengyuan Chen, Kaixiong Zhou, Daochen Zha, Xiao Huang, and Xia Hu. "*RSC: Accelerating Graph Neural Networks Training via Randomized Sparse Computations*", The 40th International Conference on Machine Learning, 2023. PDF

6. [**FBD'23**] **Zirui Liu**, Qingquan Song, Li Li, Soo-Hyun Choi, Rui Chen, and Xia Hu. "*PME: Pruning-based Multi-size Embedding for Recommender Systems*", Frontier in Big Data, 2023. PDF

7. [**NeurIPS'23**] Ruixiang Tang, Jiayi Yuan, Yiming Li, **Zirui Liu**, Rui Chen, and Xia Hu"*Setting the Trap: Capturing and Defeating Backdoor Threats in PLMs through Honeypots* ", The 37th Conference on Neural Information Processing Systems, 2023. PDF

8. [**NeurIPS'23**] Shaochen Zhong, Zaichuan You, Jiamu Zhang, Sebastian Zhao, Zachary LeClaire, **Zirui Liu**, Vipin Chaudhary, Shuai Xu, and Xia Hu"*One Less Reason for Filter Pruning: Gaining Free Adversarial Robustness with Structured Grouped Kernel Pruning*", The 37th Conference on Neural Information Processing Systems, 2023. PDF

9. [**TMLR'23**] Xiaotian Han, Zhimeng Jiang, Hongye Jin, **Zirui Liu**, Na Zou, Qifan Wang, Xia Hu. "*Retiring ΔDP: New Distribution-Level Metrics for Demographic Parity*", Transactions on Machine Learning Research, 2023. PDF

10. [**ICML'23**] Guanchu Wang, **Zirui Liu**, Zhimeng Jiang, Ninghao Liu, Na Zou, and Xia Hu. "*DIVISION: Memory Efficient Training via Dual Activation Precision*", The 40th International Conference on Machine Learning, 2023. PDF

11. [**MLSys'23**] Daochen Zha, Louis Feng, Liang Luo, Bhargav Bhushanam, **Zirui Liu**, Yusuo Hu, Jade Nie, Yuzhen Huang, Yuandong Tian, Arun Kejariwal, and Xia Hu. "*Pre-train and Search: Efficient Embedding Table Sharding with Pre-trained Neural Cost Models*", The 6th Conference on Machine Learning and Systems, 2023. PDF

12. [**ICLR'22**] **Zirui Liu**, Kaixiong Zhou, Fan Yang, Li Li, Rui Chen, and Xia Hu. "*EXACT: Scalable Graph Neural Networks Training via Extreme Activation Compression*", The 10th International Conference on Learning Representations, 2022. PDF

13. [**NeurIPS'22**] Daochen Zha, Louis Feng, Qiaoyu Tan, **Zirui Liu**, Kwei-Herng Lai, Bhargav Bhushanam, Yuandong Tian, Arun Kejariwal, and Xia Hu. " *DreamShard: Generalizable Embedding Table Placement for Recommender Systems*", The 36th Conference on Neural Information Processing Systems, 2022. PDF

14. [**Book Chatper**] Kaixiong Zhou, **Zirui Liu**, Keyu Duan, and Xia Hu. "*Graph Neural Networks: AutoML*", in **Graph Neural Networks: Foundations, Frontiers, and Applications**, 2022.

15. [**NeurIPS'22**] Keyu Duan, **Zirui Liu**, Wenqing Zheng, Peihao Wang, Kaixiong Zhou, Tianlong Chen, Zhangyang Wang and Xia Hu. "*A Comprehensive Study on Large-Scale Graph Training: Benchmarking and Rethinking*", The 36th Conference on Neural Information Processing Systems, 2022. PDF

16. [**IJCAI'22**] Kaixiong Zhou, **Zirui Liu**, Rui Chen, Li Li, Soo-Hyun Choi, and Xia Hu. "*Table2Graph: Transforming Tabular Data to Unified Weighted Graph*", The 31st International Joint Conference on Artificial Intelligence, 2022. PDF

17. [**ICLR'22**] Zhimeng Jiang, Kaixiong Zhou, **Zirui Liu**, Li Li, Rui Chen, Soo-Hyun Choi, and Xia Hu. "*An Information Fusion Approach to Learning with Instance-Dependent Label Noise*", The 10th International Conference on Learning Representations, 2022. PDF

18. [**WWW'21**] Ruixiang Tang, Mengnan Du, Yuening Li, **Zirui Liu** and Xia Hu. "*Mitigating Gender Bias in Captioning Systems*", The 32nd Web Conference, 2021. PDF

19. [**ICCV'21**] **Zirui Liu**, Haifeng Jin, Ting-Hsiang Wang, Kaixiong Zhou and Xia Hu. "*DivAug: Plug-in Automated Data Augmentation with Explicit Diversity Maximization*", The 18th International Conference on Computer Vision, 2021. PDF

20. [**NeurIPS'21**] **Zirui Liu**, Qingquan Song, Kaixiong Zhou, Ting-Hsiang Wang, Ying Shan, and Xia Hu. "*Detecting Interactions from Neural Networks via Topological Analysis*", The 35th Conference on Neural Information Processing Systems, 2021. PDF

21. [**RecSys'21**] Ting-Hsiang Wang , Qingquan Song, Xiaotian Han, **Zirui Liu**, Haifeng Jin, and Xia Hu. "*AutoRec: An Automated Recommender System*", The 15th ACM Conference on Recommender Systems (Demo track), 2021. PDF

22. [**Nano Energy'19**] Zhigang Qu, Liqun Wu, Bin Yue, Yang An, **Zirui Liu**, Pengfei Zhao, Jikui Luo, Yuedong Xie, Ying Liu, Qiuyu Wang, Zhiyuan Wang, Rui Dai, and Wuliang Yin. "*Eccentric triboelectric nanosensor for monitoring mechanical movements*", Nano Energy, 2019. PDF

23. [**TIE'19**] Ming Yang, Na Chai, **Zirui Liu**, Boyang Ren, and Dianguo Xu. "*Motor speed signature analysis for local bearing fault detection with noise cancellation based on improved drive algorithm*", IEEE Transactions on Industrial Electronics, 2019. PDF

24. [**Energies'19**] Ming Yang, **Zirui Liu**, Jiang Long, Wanying Qu, and Dianguo Xu. "*An algorithm for online inertia identification and load torque observation via adaptive Kalman observer-recursive least squares*", Energies, 2019. PDF

25. [**ITEC Asia-Pacific'17**] **Zirui Liu**, Ming Yang, Jiang Long, Dianguo Xu. "*Speed measurement error reduction via adaptive strong tracking Kalman filters*", IEEE Transportation Electrification Conference and Expo, Asia-Pacific, 2017. PDF

26. [**ITEC Asia-Pacific'17**] Yingqiang Li, Ming Yang, Jiang Long, **Zirui Liu**, and Dianguo Xu. "*Current sensorless predictive control based on extended Kalman filter for PMSM drives*", IEEE Transportation Electrification Conference and Expo, Asia-Pacific, 2017. PDF

27. [**IECON'17**] **Zirui Liu**, Ming Yang, and Dianguo Xu. "*A novel algorithm for on-line inertia identification via adaptive recursive least squares*", The 43rd Annual Conference of the IEEE Industrial Electronics Society, 2017. PDF

Recognition & Awards

- Rice Graduate Fellowship   2021
- The GMCC's Scholarship (Corporation Scholarship)   2015
- The MOONS' Scholarship (Corporation Scholarship)   2013
- First Prize of Excellent Student Scholarship of HIT   2012, 2013

Involved Research Grants

NSF Medium: Towards Effective Detection and Mitigation for Shortcut Learning: A Data Modeling Framework

$1200,000, Lead PI: Xia (Ben) Hu,   *Jun. 2022*

Led the research proposal writing on the research objectives of "Shortcut Detection via A Data-centric Framework", which is directly based on my NeurIPS paper.

NSF FAI: Towards Fairness in Deep Neural Networks with Learning Interpretation

$900,000, Lead PI: Xia (Ben) Hu,   *Sep. 2021*

Led the research proposal writing on the research objectives of "Enhancing Fairness via Feature Interpretation", which is directly based on my WWW paper.

Visa Research Award: Scalable Representation Learning Frameworks To Model Sophisticated Interactions In Large-Scale Graph And Tabular Data

$100,000, Lead PI: Xia (Ben) Hu,   *Jun. 2022*

Led the research proposal writing on the research objectives of "Very Large-Scale Graph Training", which is directly based on my ICLR & TMLR paper.

TEACHING

- Guest Lecturer, COMP 640: Graduate Seminar In Machine Learning, Rice University   2023

- Guest Lecturer, COMP 631: Introduction to Information Retrieval, Rice University   2023

- Teaching Assistant, COMP 631: Introduction to Information Retrieval, Rice University   2022, 2023

SELECTED MENTORSHIP

- **Keyu Duan**   **Fall 2021 – Present**
  Ph.D. Student, National University of Singapore   *Large-Scale Machine Learning, Neurips 2022*

- **Guanchu Wang**   **Fall 2021 – Present**
  Ph.D. Student, Rice University   *Memory-Efficient Machine Learning, ICML 2023*

- **Shengyuan Chen**   **Summer 2022 – Present**
  Ph.D. Student, The Hong Kong Polytechnic University   *Efficient Machine Learning, ICML 2023*

- **Henry Zhong**   **Fall 2022 – Present**
  Ph.D. Student, Rice University   *Efficient Machine Learning, NeurIPS 2023*

- **Jiayi Yuan**   **Fall 2022 – Present**
  Ph.D. Student, Rice University   *Machine Learning for Healthcare, AMIA*

- **Khushbu Pahwa**   **Fall 2023 – Present**
  Ph.D. Student, Rice University   *Machine Learning for Healthcare, ongoing*

- **Duy Le**   **Fall 2023 – Present**
  Undergraduate, Case Western Reserve University   *Large-Scale Machine Learning, SDM'24 in submission*

SERVICE

- **Program Committee Member**: NeurIPS, ICML, ICLR, KDD, ICCV

- **Journal Reviewers**: TPAMI, TKDD, TMLR