Jiayi Yuan

Website: `https://jy-yuan.github.io/` ◇ Email: jy101@rice.edu

EDUCATION

Rice University Houston, TX
Ph.D. in Computer Science (Advisor: Dr. Xia "Ben" Hu) *Aug. 2022 - Present*

Tsinghua University Beijing, China
B.Eng. in Computer Science *Aug. 2017 - Jul. 2021*

RESEARCH INTERESTS

Efficient and Trustworthy Machine Learning
Natural Language Processing (LLMs), Computer Vision, Health Informatics

PUBLICATION

** denotes equal contributions.*

Conference Publications

[C1] "Setting the Trap: Capturing and Defeating Backdoors in Pretrained Language Models through Honeypots", ***J. Yuan**, *R. Tang, Y. Li, Z. Liu, R. Chen, X. Hu. *In The 37th Conference on Neural Information Processing Systems (NeurIPS), 2023*

[C2] "Large language models for healthcare data augmentation: An example on patient-trial matching", **J. Yuan**, R. Tang, X. Jiang, X. Hu. *Best Student Paper. In AMIA Annual Symposium Proceedings (AMIA), 2023*

[C3] "Towards Fair Patient-Trial Matching via Patient-Criterion Level Fairness Constraint", C. Chang, **J. Yuan**, S. Ding, Q. Tan, K. Zhang, X. Jiang, X. Hu, N. Zou. *In AMIA Annual Symposium Proceedings (AMIA), 2023*

[C4] "Can Attention Be Used to Explain EHR-Based Mortality Prediction Tasks: A Case Study on Hemorrhagic Stroke", Q. Feng, **J. Yuan**, F.B. Emdad, K. Hanna, X. Hu, Z. He. *In the 14th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (ACM-BCB), 2023*

[C5] "NetBooster: Empowering Tiny Deep Learning By Standing on the Shoulders of Deep Giants", Z. Yu, Y. Fu, **J. Yuan**, H. You, Y. Lin. *In Proceedings of the 60th ACM/IEEE Design Automation Conference (DAC), 2023*

[C6] "Robust Tickets Can Transfer Better: Drawing More Transferable Subnetworks in Transfer Learning", Y. Fu, Y. Yuan, S. Wu, **J. Yuan**, Y. Lin. *In Proceedings of the 60th ACM/IEEE Design Automation Conference (DAC), 2023*

[C7] "Gen-NeRF: Efficient and Generalizable Neural Radiance Fields via Algorithm-Hardware Co-Design", *Y. Fu, *Z. Ye, **J. Yuan**, S. Zhang, S. Li, H. You, Y. Lin. *In the 50th IEEE/ACM International Symposium on Computer Architecture (ISCA), 2023*

[C8] "ERSAM: Neural Architecture Search for Energy-Efficient and Real-Time Social Ambiance Measurement", ***J. Yuan**, *C. Li, *W. Chen, Y. Lin, A. Sabharwal. *In ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023*

[C9] "DepthShrinker: A New Compression Paradigm Towards Boosting Real-Hardware Efficiency of Compact Neural Networks", Y. Fu, H. Yang, **J. Yuan**, M. Li, C. Wan, R. Krishnamoorthi, V. Chandra, Y. Lin. *In Thirty-ninth International Conference on Machine Learning (ICML), 2022*

[C10] "EyeCoD: Eye Tracking System Acceleration via FlatCam-Based Algorithm and Accelerator Co-Design", *H. You, *Y. Zhao, *Z. Yu, *C. Wan, Y. Fu, **J. Yuan**, S. Wu, S. Zhang, Y. Zhang, C. Li, V. Boominathan, A. Veeraraghavan, Z. Li, Y. Lin. *IEEE Micro Top Pick. In the 49th IEEE/ACM International Symposium on Computer Architecture (ISCA), 2022*

Preprints

[P1] "LoRA-as-an-Attack! Piercing LLM Safety Under The Share-and-Play Scenario", H. Liu, Z. Liu, R. Tang, **J. Yuan**, S. Zhong, YN. Chuang, L. Li, R. Chen, X. Hu

[P2] "KIVI: A Tuning-Free Asymmetric 2bit Quantization for KV Cache", *__J. Yuan__, *Z. Liu, H. Jin, S. Zhong, Z. Xu, V. Braverman, B. Chen, X. Hu

[P3] "Fair Patient-Trial Matching for Underrepresented Groups", C. Chang, S. Ding, **J. Yuan**, K. Zhang, X. Jiang, X. Hu, N. Zou

EXPERIENCE

Rice University Houston, TX
Graduate Research Assistant *Sep. 2022 - Present*

- Working on **Large Language Models (LLMs)**: efficient and trustworthy finetuning and inference. [P2]
- Designed a defender algorithm against natural language backdoor attacks. [C1]
- Worked on several projects regarding health informatics. [C2] [C3] [C4] [P3]

Rice University Houston, TX
Research Assistant *Aug. 2021 - Aug. 2022*

- Proposed re-parameterization-based efficient training and inference algorithms. [C5] [C9]
- Proposed a NAS pipeline for real-time social ambiance measurement. [C8]
- Worked on several projects regarding machine learning algorithms and systems co-design. [C7] [C10]
- Took part in efficient computer vision challenges: LPCVC-UAV, DAC-SDC.

Baidu Inc. Beijing, China
Research Engineer Intern *Dec. 2020 - Jul. 2021*

- Worked in Content Technology Architecture Group and took charge of processing large-scale data streams.
- Developed and optimized fingerprinting algorithms on massive real-world data. Focused on the image and video deduplication problems in both industry and academia.

Tsinghua University Beijing, China
Research Assistant *Jan. 2019 - May. 2021*

- Designed a diagnosis system for solar panel: computer vision for automated defect detection in the industry.
- Used generative models to improve Deepfake detection (forgery detection of face images).
- Built an efficient and highly scalable distributed approximate graph mining system. [Code]

HONORS AND AWARDS

SDM'24 Doctoral Forum Travel Award, ***by SIAM***	*Mar. 2024*
NeurIPS 2023 Scholar Award, ***by NeurIPS***	*Nov. 2023*
AMIA Best Student Paper, ***by AMIA***	*Nov. 2023*
AMIA 2023 KDDM Student Innovation Award, ***by AMIA***	*Oct. 2023*
IEEE Micro Top Picks, ***by ACM***	*Jul. 2023*