Yu Yang

2080 Duke University Road, Durham, NC 27708
(+1) 9196386909 | yu.yang@duke.edu | Google Scholar | GitHub

Education

Duke University — Durham, NC
Ph.D. in Electrical and Computer Engineering — *Aug. 2023 – May. 2028 (expected)*

Tongji University — Shanghai, China
B.Eng. of Software Engineering — *Sept. 2018 – Jul. 2022*

Research interests

My primary research interests focus on reinforcement learning (RL), with a specific emphasis on off-dynamics RL, RL-driven applications in healthcare, and the incorporation of RL techniques into foundation models. I am also exploring the use of RL to improve large language models (LLMs), particularly in enhancing their reasoning capabilities and alignment with human.

Publications

Yu Yang, Pan Xu. Pre-trained Language Models Improve the Few-shot Prompt Ability of Decision Transformer. *TMLR 2025.*

Yu Yang*, Jiafeng Song*, Zhishuai Liu, Henry P Foote, Rishikesan Kamaleswaran, Pan Xu. An Interactive Framework for Generating Clinical Data with Human Feedback. *BHI 2025.*

Jingxin Mao, **Yu Yang**, Zhiwei Wei, Yanlong Bi, Rongqing Zhang. D2Fed: Federated Semi-Supervised Learning With Dual-Role Additive Local Training and Dual-Perspective Global Aggregation. *TNNLS 2025.*

Tianyuan Jin, **Yu Yang**, Jing Tang, Xiaokui Xiao, Pan Xu. Optimal batched best arm identification. *NeurIPS 2024.*

Haque Ishfaq, Yixin Tan, **Yu Yang**, Qingfeng Lan, Jianfeng Lu, A Rupam Mahmood, Doina Precup, Pan Xu. More Efficient Randomized Exploration for Reinforcement Learning via Approximate Sampling. *RLC 2024.*

Yinan Huang, William Lu, Joshua Robinson, **Yu Yang**, Muhan Zhang, Stefanie Jegelka, Pan Li. On the stability of expressive positional encodings for graphs. *ICLR 2024.*

Haolong Li, Zizheng Zhong, Wei Guan, Chenghao Du, **Yu Yang**, Yuxiang Wei, Chen Ye. Generative character inpainting guided by structural information. *The Visual Computer 2021.*

Preprints

Weixin Wang*, **Yu Yang***, Pan Xu. Diffusion Posterior Sampling for Nonlinear Contextual Bandits? *Under Review.*

Zhishuai Liu, **Yu Yang**, Ruhan Wang, Pan Xu, Dongruo Zhou. How to Provably Improve Return Conditioned Supervised Learning? *NeurIPS 2025 Workshop on ARLET.*

Yihong Guo, **Yu Yang**, Pan Xu, Anqi Liu. MOBODY: Model-Based Off-Dynamics Offline Reinforcement Learning. *NeurIPS 2025 Workshop on ARLET.*

Ruhan Wang*, **Yu Yang***, Zhishuai Liu, Dongruo Zhou, Pan Xu. Return Augmented Decision Transformer for Off-Dynamics Reinforcement Learning. *NeurIPS 2025 Workshop on Reliable ML.*

Research Experience

Transformers in Reinforcement Learning. — Sept. 2023 – Present
Advisor: Prof. Pan Xu — *Duke University*

- Leveraged pre-trained language models and prompt regularization to enhance few-shot generalization in Meta-RL environments. (NeurIPS 2024 Workshop on AFM)
- Designed and conducted experiments for return-aligned decision transformer addressing offline off-dynamics RL problems. (NeurIPS 2025 Workshop on Reliable ML)
- Designed and conducted experiments on the improvement of return-conditioned reinforcement learning via quantile or expectile regression on D4RL. (NeurIPS 2025 Workshop on ARLET)

Off-dynamics Reinforcement Learning. Apr. 2024 – Present
Advisor: Prof. Pan Xu *Duke University*

- Conducted experiments on model-based methods using the source domain data to assist training the target dynamics transition model for off-dynamics RL. (NeurIPS 2025 Workshop on ARLET)
- Extended the offline off-dynamics RL to the decision transformer via augmenting the return. (see also NeurIPS 2025 Workshop on Reliable ML)
- Leveraged the target state-action pair to cluster and filter out the source dataset which are similar to the target domain. (Ongoing Project)

Reinforcement Learning for Efficient Exploration Apr. 2024 – Present
Advisor: Prof. Pan Xu *Duke University*

- Conducted experiments on leveraging an approximate posterior-sampling method for RL that reduces computational cost and maintains strong exploration efficiency in deep RL on Atari games. (RLC 2024)
- Designed and conducted experiments leveraging a diffusion prior for Thompson Sampling during the inverse stage, solving multi-task non-linear contextual bandits. (Under Review)

Reinforcement Learning for Healthcare Applications Jun. 2024 – Present
Advisor: Prof. Pan Xu *Duke University*

- Developed offline RL methods for learning clinical decision policies in VV-ECMO settings, including pipeline design and off-policy evaluation. (Under Review.)
- Leveraged the partially observed conditional generative model for evaluating the RL policy learned by the offline RL algorithm. (BHI 2025)

Work Experience

Shanghai AI Lab | Research Intern Apr. 2022 – Jul.2022

- Conducted a literature survey on recent advances in out-of-distribution(OOD) detection.
- Reproduced state-of-the-art out-of-distribution detection models for image classification and object detection tasks.
- Designed an energy-based OOD detector leveraging virtual outliers, improving robustness on CIFAR and ImageNet benchmarks.

Ant Group, Alipay Information Technology Co., Ltd. | Engineer Intern Jun. 2021 – Sept.2021

- Learned the Middleware such as Zdal and Drm and the Sofaboot framework developed by AntGroup.
- Updated the iteration using the components learned about the backend at the Risk Mesh group.
- Managed and developed the documents about the business risk and the iterations submitted.

Academic Service-Reviewer

Annual Conference on Neural Information Processing Systems (NeurIPS)	2024, 2025
International Conference on Machine Learning (ICML)	2024, 2025
International Conference on Learning Representations (ICLR)	2024, 2025(Notable)
International Conference on Artificial Intelligence and Statistics (AISTATS)	2024, 2025
Association for the Advancement of Artificial Intelligence (AAAI)	2024

Notable Awards

Outstanding Student Scholarship, Tongji University	Nov. 2020, Nov. 2021
National Encouragement Scholarship, Tongji University	Nov. 2019, Nov. 2020, Nov.2021

Technical Skills

Programming Languages: Python, Java, C, C++, C#, JavaScript, MATLAB, Latex, SQL

Frameworks and Platforms: Pytorch, PyTorch Lightning, Jax, Docker, Slurm, verl