Canzhe Zhao Ph.D.

canzhezhao@sjtu.edu.cn
https://canvolcano.github.io/

Education

Shanghai Jiao Tong University — Shanghai, China
Ph.D. in Computer Science and Engineering — 2021–2026
▷ Advisor: Prof. Shuai Li

Shandong University — Jinan, China
B.Eng. in Software Engineering — 2017–2021
▷ GPA: 4.73/5.0, ranking: 2/342.

Professional Experience

Department of Mathematics, National University of Singapore — Singapore
Research Fellow (Postdoctoral Researcher) — Jul. 2026–Present
▷ Advisor: Prof. Vincent Y. F. Tan

Internships

Department of Mathematics, National University of Singapore — Singapore
Host(s): Prof. Vincent Y. F. Tan — Jun. 2025–Dec. 2025

RIKEN Center for Advanced Intelligence Project (AIP) — Tokyo, Japan
Host(s): Prof. Shinji Ito — Sept. 2024–Jan. 2025

Tencent AI Lab — Shenzhen, China
Host(s): Weiming Liu, Haobo Fu — Jun. 2023–Sept. 2023

School of Data Science, The Chinese University of Hong Kong (Shenzhen) — Shenzhen, China
Host(s): Prof. Baoxiang Wang — Feb. 2023–May 2023

Preprints

1. Best-of-Both-Worlds FTRL for Linear Bandits: The Interplay of Information Acquisition, Corruption, and Geometry.
Canzhe Zhao, Vincent Y. F. Tan.
Under review, 2026.

Journal Articles

1. Heavy-Tailed Linear Bandits: Adversarial Robustness, Best-of-Both-Worlds, and Beyond.
Canzhe Zhao*, Shinji Ito*, Shuai Li.
JMLR, Forthcoming.

Conference Papers

1. Improved Sample Complexity for Markov Games via Variance-Aware Bandit Learning.
Hanbin Zhou*, **Canzhe Zhao***, Shuai Li.
NeurIPS, 2026.
2. No Free Best-of-Both-Worlds Learning in Repeated Bilateral Trade.
Yutian Cheng, **Canzhe Zhao**, Jingye Zhao, Shuai Li.
NeurIPS, 2026.
3. Policy Optimization for Adversarial Linear Mixture MDPs with Unknown Transitions and Bandit Feedback.
Yutian Cheng, **Canzhe Zhao**, Shuai Li.
UAI, 2026.
4. Learning Imperfect Information Extensive-form Games with Last-iterate Convergence under Bandit Feedback.
Canzhe Zhao, Yutian Cheng, Jing Dong, Baoxiang Wang, Shuai Li.
ICML, 2025.
5. Towards Provably Efficient Learning of Imperfect Information Extensive-Form Games with Linear Function Approximation.
Canzhe Zhao, Shuze Chen, Weiming Liu, Haobo Fu, Qiang Fu, Shuai Li.
UAI, 2025.
6. Logarithmic Regret for Linear Markov Decision Processes with Adversarial Corruptions.
Canzhe Zhao, Xiangcheng Zhang, Baoxiang Wang, Shuai Li.
AAAI (Oral), 2025.

7. Learning Adversarial Low-rank Markov Decision Processes with Unknown Transition and Full-information Feedback.
   Canzhe Zhao, Ruofeng Yang, Baoxiang Wang, Xuezhou Zhang, Shuai Li.
   NeurIPS, 2023.
8. Best-of-Three-Worlds Analysis for Linear Bandits with Follow-the-Regularized-Leader Algorithm.
   Fang Kong, **Canzhe Zhao**, Shuai Li.
   COLT, 2023.
9. DPMAC: Differentially Private Communication for Cooperative Multi-Agent Reinforcement Learning.
   Canzhe Zhao*, Yanjie Ze*, Jing Dong, Baoxiang Wang, Shuai Li.
   IJCAI, 2023.
10. Learning Adversarial Linear Mixture Markov Decision Processes with Bandit Feedback and Unknown Transition.
    Canzhe Zhao, Ruofeng Yang, Baoxiang Wang, Shuai Li.
    ICLR, 2023.
11. Clustering of Conversational Bandits with Posterior Sampling for User Preference Learning and Elicitation.
    Qizhi Li*, **Canzhe Zhao***, Tong Yu, Junda Wu, Shuai Li.
    UMUAI, 2023.
12. Differentially Private Temporal Difference Learning with Stochastic Nonconvex-Strongly-Concave Optimization.
    Canzhe Zhao, Yanjie Ze, Jing Dong, Baoxiang Wang, Shuai Li.
    WSDM, 2023.
13. Knowledge-aware Conversational Preference Elicitation with Bandit Feedback.
    Canzhe Zhao, Tong Yu, Zhihui Xie, Shuai Li.
    WWW, 2022.
14. Simultaneously Learning Stochastic and Adversarial Bandits under the Position-based Model.
    Cheng Chen, **Canzhe Zhao**, Shuai Li.
    AAAI, 2022.
15. Clustering of Conversational Bandits for User Preference Learning and Elicitation.
    Junda Wu*, **Canzhe Zhao***, Tong Yu, Jingyang Li, Shuai Li.
    CIKM, 2021.
16. Comparison-based Conversational Recommender System with Relative Bandit Feedback.
    Zhihui Xie, Tong Yu, **Canzhe Zhao**, Shuai Li.
    SIGIR, 2021.

Honors and Awards

Top Reviewer, NeurIPS 2026 — Oct. 2026
Shanghai Outstanding Ph.D. Graduate — May 2026
Travel Award, Maoshisong Statistics Education Doctoral Forum and 2026 Annual Meeting of the National Young Statisticians Chapter — Apr. 2026
National Scholarship, Ministry of Education, People's Republic of China — Oct. 2025
Top Reviewer, NeurIPS 2025 — Oct. 2025
Huatai Securities Scholarship, Shanghai Jiao Tong University — Dec. 2022
The First Prize, The Third Doctoral Academic Forum of John Hopcroft Center — Dec. 2022
Silver Medal (**Top** 6%), Kaggle Santa 2020 – The Candy Cane Contest — Jan. 2021
Bronze Medal, The 2019 ACM-ICPC Asia Regional Contest — Oct. 2019
Weichai Power Scholarship, Shandong University — Oct. 2019
Silver Medal, The 2019 ACM-ICPC China Shandong Provincial Contest — May 2019
Huawei Scholarship, Shandong University — Oct. 2018

Talks

Tutorial for **Theoretical Foundations for Markov Games**, AAMAS 2025 — May 2025

Academic Services

Reviewer for

NeurIPS (2021–)	ICML (2021–)	ICLR (2023–)
AISTATS (2023–)	UAI (2021–2023)	ALT (2024–)
AAMAS (2021–2022)	JSAIT (2021)	TNNLS (2023)

Teaching Experience

CS 3317: Artificial Intelligence — Shanghai Jiao Tong University
Teaching Assistant — Spring 2025

AI 3601: Reinforcement Learning — Shanghai Jiao Tong University
Teaching Assistant — Spring 2023

CS 2601: Linear and Convex Optimization — Shanghai Jiao Tong University
Teaching Assistant — Fall 2022

CS 410: Artificial Intelligence — Shanghai Jiao Tong University
Teaching Assistant — Fall 2021