

# Chenjie Mao

[cmiao@wustl.edu](mailto:cmiao@wustl.edu)

## ABOUT ME

---

I am a PhD student at Washington University in St. Louis, advised by Professor Chongjie Zhang. My research interests encompass a broad range of topics in learning theory, including reinforcement learning, online learning, and deep learning theory. Recently, I have focused primarily on the statistical aspects of (offline) reinforcement learning, especially its non-asymptotic behaviors and extensions to reinforcement learning with human feedback (RLHF).

## EDUCATION

---

**Huazhong University of Science and Technology** 2020 – 2024  
Bachelor of Engineering, School of Computer Science and Technology  
**Huazhong University of Science and Technology** 2025 – Present  
Phd Student, Computer Science and Engineering

## WORK EXPERIENCE

---

**Shanghai Artificial Intelligence Laboratory** 2023.7 – 2025.6  
*Research Assistant*

- Investigated the statistical aspects of offline reinforcement learning and related areas.
- Explored potential alignments between RL and RLHF.

## PREPRINTS & PUBLICATIONS

---

### 1. Offline Reinforcement Learning with Additional Covering Distributions

Chenjie Mao<sup>1</sup>

*TMLR (Transactions on Machine Learning Research)*

- Proposed an algorithm with statistical guarantees for offline RL, utilizing general function approximation under single coverage and realizability-type assumptions, based on an additional covering distribution.

### 2. On the Role of General Function Approximation in Offline Reinforcement Learning

Chenjie Mao, Qiaosheng Zhang, Zhen Wang, Xuelong Li

*ICLR 2024 Spotlight, Top 5%*

- Investigated the statistical limitations of offline reinforcement learning with general function approximation and established some lower bounds.

### 3. Rethinking the Hardness of PbRL: A Provable General Regret Bound

Chenjie Mao, Yi Fan, Ning Zhang, Chongjie Zhang

*International Conference on Machine Learning (ICML), 2026*

- Introduced RTPQ, a simple and provably efficient algorithm that achieves near-optimal  $\sqrt{T}$  regret in preference-based reinforcement learning with trajectory-level feedback.

## INTERESTS

---

I have broad interests in topics in machine learning, such as:

- Reinforcement Learning
- Generative Models

## SKILLS

---

**Proficient:** Python, PyTorch, Git, Linux, LaTeX

**Have Experience:** Shell Scripting, C/C++, Golang

---

<sup>1</sup>Completed independently as a third-year undergraduate student.