

CHENLU YE

University of Illinois Urbana-Champaign
chenluy3@illinois.edu | <https://chenluye99.github.io/>

RESEARCH INTERESTS

- Reinforcement learning from human feedback (RLHF) for aligning large language model;
- Reinforcement Learning Theory;
- Statistical Machine Learning.

EDUCATION

University of Illinois Urbana-Champaign

Ph.D. student, *Computer Science*

Advisor: Prof. Tong Zhang

Urbana, USA

2024.8 - present

The Hong Kong University of Science and Technology

MPhil, *Artificial Intelligence*

Advisor: Prof. Tong Zhang and Prof. Kani Chen, Prof. Yuan Yao

Hong Kong, China

2021.9 - 2024.8

University of Science and Technology of China

Bachelor of Science, *Statistics*

Ranking: Top 20% in Statistics.

Hefei, China

2017.9 - 2021.6

RESEARCH

Preference Learning from Human Feedback under KL-Regularization

2024.08 - present

University of Illinois Urbana-Champaign

Advisor: Prof. Tong Zhang from UIUC, Prof. Quanquan Gu from UCLA

- We studies the role of KL-regularization and data coverage for RLHF and regularized contextual bandits.
- We found that the KL-constraint leads to a better sample complexity, and the sample complexity also enjoys a better coverage dependence on the data-collector policy under an on-policy framework.

Algorithm Designs in Reinforcement Learning from Human Feedback (RLHF) *2023.11 - 2024.8*

The Hong Kong University of Science and Technology

Advisor: Prof. Tong Zhang, Prof. Nan Jiang from UIUC

- We formulated the real-world RLHF process as a reverse-KL regularized contextual bandits for preference satisfying Bradley-Terry (BT) model and a the reverse-KL regularized minimax game under general preference, respectively. We studied its theoretical property by proposing statistically efficient algorithms with finite-sample theoretical guarantee.
- We connected our theoretical findings with practical algorithms (e.g. DPO, RSO, iterative IPO), offering new tools and insights for the algorithmic design of alignment algorithms.

Corruption-Robust Reinforcement Learning with General Function Approximation *2022.9 - 2024.2*

University of California, Los Angeles, and The Hong Kong University of Science and Technology

Advisor: Prof. Tong Zhang, Prof. Quanquan Gu from UCLA

- We developed a series of corruption-robust algorithms based on uncertainty weighting for online and offline, value-based and model-based settings.
- We provided theoretical analysis for each algorithm, which enjoy an optimal regret dependence on the corruption level. We implemented the offline algorithm practically under various data-corruption scenarios, which outperforms the state-of-the-art.

Optimal Sample Selection Through Uncertainty Estimation and Its Application in Deep Learning

2020.3 - 2021.8

HKUST

Advisor: Prof. Tong Zhang, and Prof. Yuan Yao from HKUST

- We proposed a theoretically optimal and computationally efficient sample selection approach.
- We effectively applied it to deep learning and is robust to misspecification (by down-weighting highly uncertain samples).

Provably Efficient Learning in High-Dimensional Batched Bandits

2020.3-2021.7

USTC

Advisor: Prof. Zhaoran Wang from Northwestern University

- We designed an efficient algorithm for high-dimensional linear contextual bandits with batched feedback.
- Our algorithm enjoyed nearly the same regret order as the sequential case.
- We extended the algorithm to low-rank matrix bandits.

EXPERIENCE

University of California, Los Angeles:

Visiting Research Scholar advised by Prof. Quanquan Gu

2023.8 - 2023.12

The Hong Kong University of Science and Technology:

Teaching Assistant: EMIA 2020 - Cross-disciplinary Design Thinking

Fall 2022

University of Science and Technology of China:

Teaching Assistant: Mathematical Statistics

Fall 2020

Teaching Assistant: Linear Algebra

Spring 2020

HONORS AND AWARDS

Gold Prize for Outstanding Student Scholarship (1/40)

2020.9

Bronze Prize for Outstanding Student Scholarship (Top 20%)

2019.9

Bronze Prize for Outstanding Student Scholarship (Top 20%)

2018.9

SELECTED PUBLICATIONS AND PREPRINTS

(* denotes alphabetical order or equal contribution)

- [1] [Chenlu Ye*](#), Wei Xiong*, Yuheng Zhang*, Hanze Dong*, Nan Jiang, Tong Zhang, “A theoretical analysis of nash learning from human feedback under general kl-regularized preference”, [\[preprint\]](#).
- [2] Wei Xiong*, Hanze Dong*, [Chenlu Ye*](#), Han Zhong, Nan Jiang, Tong Zhang, “Iterative preference learning from human feedback: Bridging theory and practice for rlhf under kl-constraint”, [\[ICML 2024\]](#)
- [3] [Chenlu Ye*](#), Jiafan He*, Quanquan Gu, Tong Zhang, “Towards robust model-based reinforcement learning against adversarial corruption”, [\[ICML 2024\]](#).

- [4] Chenlu Ye*, Rui Yang*, Quanquan Gu and Tong Zhang, “Corruption-Robust Offline Reinforcement Learning with General Function Approximation”, [[NeurIPS 2023](#)].
- [5] Yong Lin*, Chen Liu*, Chenlu Ye*, Qing Lian, Yuan Yao and Tong Zhang, “Optimal Sample Selection Through Uncertainty Estimation and Its Application in Deep Learning”, [[Preprint](#)].
- [6] Chenlu Ye, Wei Xiong, Quanquan Gu, and Tong Zhang, “Corruption-Robust Algorithms with Uncertainty Weighting for Nonlinear Contextual Bandits and Markov Decision Processes”, [[ICML 2023](#)].
- [7] Jianqing Fan*, Zhaoran Wang*, Zhuoran Yang*, Chenlu Ye*, “Provably Efficient High-Dimensional Bandit Learning with Batched Feedbacks”, [[Preprint](#)].

PROFESSIONAL ACTIVITY

Conference Reviewer: ICML 2024.

Journal Reviewer: Machine Learning, Artificial Intelligence.