MIAO LU

School of the Gifted Young, University of Science and Technology of China
(+86)15861397762 | lumiao@mail.ustc.edu.cn | http://miaolu3.github.io

EDUCATION

University of Science and Technology of China *Hefei, China*

- School of the Gifted Young (skipped the 3rd year in high school, with summa cum laude). *Sep.2018 - Jun.2022*
- School of Mathematical Sciences, Department of Probability and Statistics. *Sep.2019 - Jun.2022*
- B.S. in Mathematics and Applied Mathematics, **Overall GPA:** 4.06/4.3, **Ranking:** **2**/140
- **Course Highlights:** *Real Analysis* (100), *Functional Analysis* (100), *(Elementary, Advanced*) Probability Theory* (95, 98), *(Applied, Advanced*) Stochastic Process* (97, 97), *Mathematical Statistics* (95), *Data Structure and Data Base* (96), *Regression Analysis* (90), *Martingale Theory and Stochastic Integration** (98), *Non-parametric Statistics** (89), *Bayesian Analysis** (92), *Optimization Algorithms** (99), *Information Theory** (93). (*: graduate-level)

HONORS AND SCHOLARSHIPS

The 41st Guo Moruo Scholarship. (Top 1%, highest honor & scholarship from USTC) *Sep.2021*
Yuanqing Yang Scholarship. (Top 2%, School of Mathematical Sciences at USTC) *Dec.2021*
National Scholarship. (Top 2.5%, highest scholarship from Ministry of Education of China) *Oct.2019, 2020*
S.-T.Yau College Student Mathematics Contests, Winning Prize. (From Tsinghua University) *Nov.2020*
Outstanding Freshman Scholarship. (From USTC awarding outstanding freshmam) *Sep.2018*

RESEARCH INTERESTS

The goal of my researches is to develop efficient data-driven decision-making algorithms and theory, enabling principled approaches via statistics and machine learning to addressing modern societal, medical, and economic challenges. Specifically, I'm interested in the following perspectives and topics:

- Demystifying the mathematical and statistical foundations of machine learning/reinforcement learning algorithms.
- Promoting robustness of machine learning methods in complex dynamic systems, with special attention to problems involving distributional robustness, confounding, and partial observability.
- Developing modern machine learning models with low computational and memorial costs.

With these goals, my researches are based across the span of statistics, optimization, and machine learning.

PUBLICATIONS

[1] Pessimism in the Face of Confounders: Provably Efficient Offline Reinforcement Learning in Partially Observable Markov Decision Processes.
Miao Lu, Yifei Min, Zhaoran Wang, Zhuoran Yang
ArXiv Preprint, Submitted to *International Conference on Learning Representation (ICLR) 2023*, Pdf.
Presented by Prof. Zhuoran Yang at INFORMS2022 and Stat.&ML Workshop at University of Michigan, Slide.

[2] Welfare Maximization in Competitive Equilibrium: RL for Markov Exchange Economy.
Zhihan Liu*, **Miao Lu***, Zhaoran Wang, Michael I. Jordan, Zhuoran Yang
Accepted to *International Conference on Machine Learning (ICML) 2022* as poster (acc. rate: 19.8%), Pdf, Code.

[3] Learning Pruning-Friendly Networks via Frank-Wolfe: One-Shot, Any-Sparsity, and No Retraining.
Miao Lu*, Xiaolong Luo*, Tianlong Chen, Wuyang Chen, Dong Liu, Zhangyang Wang
Accepted to *International Conference on Learning Representation (ICLR) 2022* as **spotlight** (top 5.2%), Pdf, Code.

[4] Pessimistic Offline Reinforcement Learning with Confounded Offline Data in Large State Space: Linear Function Approximations and Beyond.
Miao Lu, Zhaoran Wang, Zhuoran Yang, Xi Chen
In preparation for submission to *Operations Research (OR)*. Manuscript available upon request.

[5] Learning Robust Policy against Disturbance in Transition Dynamics via State-Conservative Policy Optimization.
Yufei Kuang, **Miao Lu**, Jie Wang, Qi Zhou, Bin Li, Houqiang Li
Accepted to *Association for Advancement of Artificial Intelligence (AAAI) 2022* as poster (acc. rate: 14.6%), Pdf, Code.

(Note: Authors with * contributed equally to the work.)

PROFESSIONAL EXPERIENCES

Ubiquant Investment, Research Intern (2022 Summer in Shanghai, China) *Jun.2022 - Sep.2022*

- Quantitative research intern at the AI Department of Ubiquant Investment. MENTOR: Dr. Ningshan Zhang.
- Applying modern machine learning and deep learning techniques in quantitative trading.

USTC Teaching Assistant, Differential Equation 1 (2020 Fall at USTC) *Sep.2020 - Feb.2021*

- INSTRUCTOR: Prof. Wuqing Ning (Department of Applied Mathematics@USTC).
- TEACHING CONTENTS: theory of ODE and classical PDE.

RESEARCH EXPERIENCES

Data-driven Decision Making *Sep.2021 - May.2022*

PROJECT: Reinforcement learning for Markovian exchange economy. (Results in a paper [**2**])
ADVISOR: Prof. **Zhuoran Yang** (Statistics & Data Science@Yale), Prof. **Zhaoran Wang** (IEMS@NU), Prof. **Michael I. Jordan** (EECS@UC Berkeley)

- We propose and study a bilevel economic system, known as a Markov Exchange Economy (MEE), from the viewpoint of multi-agent reinforcement learning. With a central planner and a group of agents, the system is optimized when agents achieve competitive equilibrium (CE) while planner steers the economy to social welfare maximization (SWM).
- We derived a novel metric to characterize such an optimality, based on which we design provably efficient online and offline learning algorithms for solving the economy. Our algorithm can readily incorporate general function approximation tools for handling large state spaces and achieves a $\mathcal{O}(K^{1/2})$-online regret or $\mathcal{O}(n^{-1/2})$-offline suboptimality.

Reinforcement Learning Theory *Feb.2022 - present*

PROJECT: Offline policy optimization in partially observable Markov decision processes. (Results in a paper [**1**])
ADVISOR: Prof. **Zhaoran Wang** (IEMS@NU), Prof. **Zhuoran Yang** (Statistics & Data Science@Yale)

- We propose the *first* provably efficient offline RL algorithm known as `P3O` for POMDPs with a *confounded dataset*.
- At the core of `P3O` is a coupled sequence of pessimistic confidence regions constructed via proximal causal inference, which is formulated as minimax estimation. We also derived novel theoretical analysis techniques to show the fast statistical rate for the confidence region of this kind, allowing us to prove a $\mathcal{O}(n^{-1/2})$-subobtimality rate of `P3O`.

Robust Deep Reinforcement Learning *Oct.2020 - Aug.2021*

PROJECT: Learning robust policy against disturbance in transition dynamics. (Results in a paper [**4**])
ADVISOR: Prof. **Jie Wang** (EEIS@USTC), Prof. **Houqiang Li** (EEIS@USTC)

- We propose a State-Conservative Markov Decision Process (SC-MDP) to learn robust policies against disturbance in transition dynamics without task-specific knowledge of the disturbance.
- From the theoretical perspective, in the tabular setting, we design the State-Conservative Policy Iteration algorithm (`SCPI`) to learn the correponding optimal policy, which enjoys convergence guarantees.
- To promote robustness in continuous control tasks, we further propose the State-Conservative Policy Optimization algorithm (`SCPO`) based on `SCPI`, and we efficiently implement it via a Gradient Based Regularizer (GBR). Mujoco experiments show that `SCPO` can improve robustness against disturbance in transition dynamics in various domains.

Deep Learning Model Compression *May.2021 - May.2022*

PROJECT 1: Network pruning via Stochastic Frank-Wolfe: any sparsity and no retraining. (Results in a paper [**3**])
ADVISOR: Prof. **Zhangyang (Atlas) Wang** (ECE@UT Austin), **Dong Liu** (EEIS@USTC)

- We propose Stochastic Frank-Wolfe Pruning (SFW-pruning), a one-shot unstructured deep neural network pruning algorithm. It results in consistent and competitive model performance for varying pruning ratios *without retraining*.
- We customize a meta-learning-based initialization scheme for SFW-based DNN training, leading to more consistent and competitive performance under more varying pruning ratios.
- We demonstrate the efficiency and competitiveness of SFW-pruning over various DNN models and CV datasets.

PROJECT 2: Neural tangent kernel and NN compression at initialization. (Undergraduate Research Program@USTC)
ADVISOR: Prof. **Richard Xu** (Math@HKBU), **Weiping Zhang** (Statistics@USTC)

- We first study basic knowledges of neural tangent kernel (NTK) and its application in proving the global convergence for train overparameterized neural networks via gradient descent.
- We further consider a kind of model compression scheme at initialization via matrix factorization, and we extend existing works using NTK to this setting. We also conduct experiments to verify our theoretical findings.

SKILLS AND STRENGTHS

Programming Languages and Tools. Fluent: Python, Pytorch, LaTeX, Git; Familiar: C/C++, R
Languages. Chinese (native), English (TOEFL:104/120 with Reading 27, Listening 28, Speaking 23, Writing 26)