

# Yao Liu

yao.liu.chn@gmail.com  $\diamond$  <http://www.yao-liu.com>

## RESEARCH INTERESTS

---

Reinforcement learning, interactive learning, causal inference, statistical learning, and applications in healthcare and recommendation systems

## EDUCATION EXPERIENCE

---

**Ph.D. Candidate in Computer Science, Stanford University**

*June 2021*

Dissertation: Adaptive and Efficient Batch Reinforcement Learning Algorithms

Thesis Reading Committee: Emma Brunskill (Advisor), Tengyu Ma, Stefan Wager, Benjamin Van Roy

**B.S. in Machine Intelligence, Peking University**

*July 2016*

GPA: 3.75/4.0, Rank: 1/46. Summa Cum Laude,

## WORK EXPERIENCE

---

**Research Scientist**

*TikTok*

Aug 2021 -

*Bellevue, WA*

- Improving user growth and recommendation system by RL/bandit algorithms and causal inference.
  - Lead a project about personalized push notification frequency for user growth: built up a causal bandit model that can jointly leverage both confounding data and RCT data to balance data efficiency and confounding bias, achieved 0.1% App DAU gain.
  - Improving ads recommendation by off-policy contextual bandit learning.

## PUBLICATIONS

---

### Peer-Reviewed Publications:

1. **Yao Liu**, Adith Swaminathan, Alekh Agarwal, Emma Brunskill  
Provably Good Batch Off-Policy Reinforcement Learning Without Great Exploration  
(NeurIPS'20) In Advances in Neural Information Processing Systems 33
2. **Yao Liu**, Pierre-Luc Bacon, Emma Brunskill  
Understanding the Curse of Horizon in Off-Policy Evaluation via Conditional Importance Sampling  
(ICML'20) In Proceedings of the 37th International Conference on Machine Learning
3. Omer Gottesman, Joseph Futoma, **Yao Liu**, Sonali Parbhoo, Leo Anthony Celi, Emma Brunskill, Finale Doshi-Velez  
Interpretable Off-Policy Evaluation in Reinforcement Learning by Highlighting Influential Transitions  
(ICML'20) In Proceedings of the 37th International Conference on Machine Learning
4. **Yao Liu**, Adith Swaminathan, Alekh Agarwal, Emma Brunskill  
Off-Policy Policy Gradient with State Distribution Correction  
(UAI'19, **Oral**) In Proceedings of the 35th Conference on Uncertainty in Artificial Intelligence
5. Omer Gottesman, **Yao Liu**, Scott Sussex, Emma Brunskill, Finale Doshi-Velez  
Combining Parametric and Nonparametric Models for Off-Policy Evaluation  
(ICML'19, **Oral**) In Proceedings of the 36th International Conference on Machine Learning

6. **Yao Liu**, Omer Gottesman, Aniruddh Raghu, Matthieu Komorowski, Aldo Faisal, Finale Doshi-Velez, Emma Brunskill  
Representation Balancing MDPs for Off-Policy Policy Evaluation  
(NeurIPS'18) Advances in Neural Information Processing Systems 31
7. **Yao Liu**, Zhaohan Daniel Guo, Emma Brunskill  
PAC Continuous State Online Multitask Reinforcement Learning with Identification  
(AAMAS'16) In Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems
8. Tong Lin, **Yao Liu**, Bo Wang, Liwei Wang, Hongbin Zha  
Local Orthogonality Preserving Alignment for Nonlinear Dimensionality Reduction  
Journal of Computer Science and Technology, 2016, Vol. 31, Issue 3: Page 512-524.

### Preprints and Workshop Papers:

1. **Yao Liu**, Yannis Flet-Berliac, Emma Brunskill  
Offline Policy Optimization with Eligible Actions  
In submission
2. Benjamin Petit, Loren Amdahl-Culleton, **Yao Liu**, Jimmy Smith, Pierre-Luc Bacon  
All-Action Policy Gradient Methods: A Numerical Integration Approach  
(NeurIPS'19 Workshops) arXiv preprint arXiv:1910.09093
3. Aniruddh Raghu, Omer Gottesman, **Yao Liu**, Matthieu Komorowski, Aldo Faisal, Finale Doshi-Velez, Emma Brunskill  
Behaviour Policy Estimation in Off-Policy Evaluation: Calibration Matters  
(ICML'18 Workshop) arXiv preprint arXiv:1807.01066
4. Scott Sussex, Omer Gottesman, **Yao Liu**, Susan Murphy, Emma Brunskill, Finale Doshi-Velez  
Switched Trajectories for Off-Policy Learning  
(ICML'18 Workshop)
5. **Yao Liu**, Emma Brunskill  
When Simple Exploration is Sample Efficient: Identifying Sufficient Conditions for Random Exploration to Yield PAC RL Algorithms  
(EWRL'18) Presented in the 14th European Workshop on Reinforcement Learning
6. **Yao Liu**, Philip S. Thomas, Emma Brunskill  
Model Selection for Off-Policy Policy Evaluation (Extended abstract)  
(RLDM'17) Presented in the 3rd Multidisciplinary Conference on Reinforcement Learning and Decision Making, 2017. Extended abstract.

## VISITING AND INTERNSHIPS

---

### Visiting Graduate Student

*Simons Institute for the Theory of Computing*

Sept 2020 - Dec 2020

*Berkeley, CA*

- Remote visit to the theory of reinforcement learning program.

### Research Intern

*Microsoft Research*

June 2020 - Sept 2020

*New York, NY*

- Developed provable sim-to-real transfer algorithm from latent MDP to high-dimensional observations.
- Mentors: Dipendra Misra, Miro Dudk, and Robert Schapire.

### Data Science Intern

*Livongo Health*

June 2019 - Sept 2019

*Mountain View, CA*

- Implemented offline evaluation for personalized nudge to help improve healthy behavior.

**Research Intern**  
*Microsoft Research*

June 2018 - Sept 2018  
*Redmond, WA*

- Developed provably convergent offline policy gradient algorithm with non-expert dataset.
- Mentors: Adith Swaminathan and Alekh Agarwal.

## TEACHING EXPERIENCE

---

**Stanford University**  
*Teaching Assistant*

- CS234: Reinforcement Learning, CS 229: Machine Learning

**Peking University**  
*Teaching Assistant*

- Algorithm Design and Analysis

## AWARDS

---

School of Engineering Fellowship, Stanford University	Sept 2017 - June 2018
Undergraduate Research Award, School of EECS, Peking University	Nov 2015
Lee Wai Wing Scholarship, Peking University	Sept 2015
Arawana Scholarship, Peking University	Sept 2014
Founder Scholarship, Peking University	Sept 2013
First Award, National Mathematical Olympiad in Provinces (A country-wide high school mathematics competition)	Sept 2011

## INVITED TALKS

---

Provably Good Batch Reinforcement Learning Without Great Exploration. UC Berkeley, Host: Prof Jiantao Jiao	Oct 2020
Provably Good Batch Reinforcement Learning Without Great Exploration. Statistical Machine Learning Lunch, Stanford	Oct 2020
On the Variance of Conditional Importance Sampling for Off-Policy Evaluation Causal Inference Seminar, Stanford	Oct 2019

## ACADEMIC AND COMMUNITY SERVICES

---

Reviewer for *Biometrika*, *Journal of Machine Learning Research (JMLR)*, *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, *Machine Learning*, *Artificial Intelligence*.

Reviewer for *NeurIPS* (2019-), *ICML*(2020-), *ICLR* (2019-), *AISTATS* (2020-), *UAI*(2020)

Served on the admission committee of Computer Science PhD program at Stanford: 2020, 2021.