# Shangtong Zhang

Department of Computer Science
School of Engineering and Applied Science
University of Virginia
Charlottesville, VA, 22903

Rice Hall 422
85 Engineer's Way
Charlottesville, VA, 22903
shangtong@virginia.edu

## RESEARCH INTEREST

The goal of my research is to solve sequential decision making problems in a scalable and reliable way. Currently, I focus on Reinforcement Learning (RL) as a solution method. In particular, I work on stochastic approximations for RL, theories and algorithms of RL, and applications by RL.

## ACADEMIC EMPLOYMENTS

**Assistant Professor**　　　　　　　　　　　　　　　　　　　　　　　Aug 2022 - Present
Department of Computer Science
University of Virginia, VA, United States

**Research Scientist Interns**
Microsoft Research Montreal　　　　　　　　　　　　　　　　　　Jun 2021 - Sep 2021
DeepMind London　　　　　　　　　　　　　　　　　　　　　　　Feb 2021 - Jun 2021
Microsoft Research Montreal　　　　　　　　　　　　　　　　　　Jun 2020 - Aug 2020

## EDUCATION

**Doctor of Philosophy**, Computer Science　　　　　　　　　　　　Oct 2018 - Jul 2022
University of Oxford, Oxford, United Kingdom
Advisor: Prof. Shimon Whiteson
Thesis: Breaking the Deadly Triad in Reinforcement Learning

**Master of Science**, Computer Science　　　　　　　　　　　　　Sep 2016 - Jul 2018
University of Alberta, Edmonton, Canada
Advisor: Prof. Richard S. Sutton
Thesis: Learning with Artificial Neural Networks

**Bachelor of Science**, Computing Science　　　　　　　　　　　　Sep 2012 - Jul 2016
Fudan University, Shanghai, China

## HONORS

NeurIPS Top Area Chair, 2025
Cisco Faculty Research Award, 2025
NSF CAREER Award, 2025
QuantCo Spotlight (Best Paper) Award at the ICML In-Context Learning Workshop, 2024
Rising Stars in AI by KAUST AI Initiative, 2024
AAAI New Faculty Highlights, 2023
IFAAMAS Victor Lesser Dissertation Award (Runner-Up), 2022

Alf Weaver Junior Faculty Fellowship, UVA, 2022 - 2027
ICLR Outstanding Reviewer, 2021
NeurIPS Reviewer Award, 2020
ICML Reviewer Award, 2020
AAMAS Best Paper Award, 2020
Light Senior Scholarship, St Catherine's College, University of Oxford, 2020
EPSRC Studentship, University of Oxford, 2018 - 2022
EMC Scholarship, Fudan University, 2014

## PUBLICATIONS

Advisees of SZ are underlined; $^*$ indicates equal contribution; $^\dagger$ indicates equal advising.

### Preprints

11. *MathlibLemma: Folklore Lemma Generation and Benchmark for Formal Mathematics.*
    Xinyu Liu, Zixuan Xie, Amir Moeini, Claire Chen, Shuze Liu, Yu Meng, Aidong Zhang,
    **Shangtong Zhang**.
    ArXiv Preprint 2026.

10. *Multi-agent DRL-based Lane Change Decision Model for Cooperative Planning in Mixed Traffic.*
    Zeyu Mu, **Shangtong Zhang**, B. Brian Park.
    ArXiv Preprint 2026.

9. *Prompt-Driven Domain Adaptation for End-to-End Autonomous Driving via In-Context RL.*
   Aleesha Khurram, Amir Moeini, **Shangtong Zhang**, Rohan Chandra.
   ArXiv Preprint 2025.

8. *Towards Formalizing Reinforcement Learning Theory.*
   **Shangtong Zhang**.
   ArXiv Preprint 2025.

7. *Extensions of Robbins-Siegmund Theorem with Applications in Reinforcement Learning.*
   Xinyu Liu, Zixuan Xie, **Shangtong Zhang**.
   ArXiv Preprint 2025.

6. *Safe In-Context Reinforcement Learning.*
   Amir Moeini$^*$, Minjae Kwon$^*$, Alper Kamil Bozkurt, Yuichi Motai, Rohan Chandra, Lu Feng,
   **Shangtong Zhang**.
   ArXiv Preprint 2025.

5. *Experience Replay Addresses Loss of Plasticity in Continual Learning.*
   Jiuqi Wang, Rohan Chandra, **Shangtong Zhang**.
   ArXiv Preprint, 2025.

4. *A Survey of In-Context Reinforcement Learning.*
   Amir Moeini, Jiuqi Wang, Jakob Beck, Ethan Blaser, Shimon Whiteson, Rohan Chandra,
   **Shangtong Zhang**.
   ArXiv Preprint, 2025.

3. *Almost Sure Convergence Rates and Concentration of Stochastic Approximation and Reinforcement Learning with Markovian Noise.*
   Xiaochi Qian$^*$, Xinyu Liu$^*$, Zixuan Xie$^*$, **Shangtong Zhang**.
   ArXiv Preprint, 2024.

2. *Almost Sure Convergence of Linear Temporal Difference Learning with Arbitrary Features.*
   Jiuqi Wang, **Shangtong Zhang**.
   ArXiv Preprint, 2024.

1. *AlphaStar Unplugged: Large Scale Offline Reinforcement Learning.*
   Michael Mathieu*, Sherjil Ozair*, Srivatsan Srinivasan*, Caglar Gulcehre*, **Shangtong Zhang**\*,
   Ray Jiang*, Tom Le Paine*, Richard Powell, Konrad Zolna, Julian Schrittwieser, David Choi, Petko
   Georgiev, Daniel Kenji Toyama, Aja Huang, Roman Ring, Igor Babuschkin, Timo Ewalds, Mahyar
   Bordbar, Sarah Henderson, Sergio Gomez Colmenarejo, Aaron van den Oord, Wojciech M. Czarnecki,
   Nando de Freitas, Oriol Vinyals.
   ArXiv Preprint, 2023

## Invited Articles

1. *A New Challenge in Policy Evaluation.*
   **Shangtong Zhang**.
   AAAI Conference on Artificial Intelligence (**AAAI**), 2023
   New Faculty Highlights Program.

## Refereed Journals

4. *The ODE Method for Stochastic Approximation and Reinforcement Learning with Markovian Noise.*
   Shuze Liu, Shuhang Chen, **Shangtong Zhang**.
   Journal of Machine Learning Research (**JMLR**), 2025.

3. *Global Optimality and Finite Sample Analysis of Softmax Off-Policy Actor Critic under State
   Distribution Mismatch.*
   **Shangtong Zhang**, Remi Tachet des Combes[†], Romain Laroche[†].
   Journal of Machine Learning Research (**JMLR**), 2022.

2. *Truncated Emphatic Temporal Difference Methods for Prediction and Control.*
   **Shangtong Zhang**, Shimon Whiteson.
   Journal of Machine Learning Research (**JMLR**), 2022.

1. *MLPack 3: A Fast, Flexible Machine Learning Library.*
   Ryan Curtin, Marcus Edel, Mikhail Lozhnikov, Yannis Mentekidis, Sumedh Ghaisas, **Shangtong
   Zhang**
   Journal of Open Source Software (**JOSS**), 2018.

## Refereed Conference Papers

34. *Reward Is Enough: LLMs Are In-Context Reinforcement Learners.*
    Kefan Song*, Amir Moeini*, Peng Wang, Lei Gong, Rohan Chandra, **Shangtong Zhang**[†],
    Yanjun Qi[†].
    International Conference on Learning Representations (**ICLR**), 2026.
    Acceptance rate: 28%

33. *Adaptive Policy Selection and Fine-Tuning under Interaction Budgets for Offline-to-Online
    Reinforcement Learning*
    Alper Bozkurt, Xiaoan Xu, **Shangtong Zhang**, Miroslav Pajic, Yuichi Motai
    Annual Learning for Dynamics & Control Conference (**L4DC**), 2026
    Acceptance rate: not released

32. *Almost Sure Convergence of Differential Temporal Difference Learning for Average Reward Markov
    Decision Processes.*
    Ethan Blaser, Jiuqi Wang, **Shangtong Zhang**.
    International Conference on Artificial Intelligence and Statistics (**AISTATS**), 2026
    Acceptance rate: 28%

31. *Asymptotic and Finite Sample Analysis of Nonexpansive Stochastic Approximations with Markovian Noise.*
    Ethan Blaser, **Shangtong Zhang**.
    AAAI Conference on Artificial Intelligence (**AAAI**), 2026
    Acceptance rate: 17.6%

30. *GameChat: Multi-LLM Dialogue for Safe, Agile, and Socially Optimal Multi-Agent Navigation in Constrained Environments.*
    Vagul Mahadevan, **Shangtong Zhang**, Rohan Chandra.
    IEEE International Symposium on Multi-Robot and Multi-Agent Systems **(MRS)**, 2025
    Acceptance rate: 47.8%

29. *Towards Provable Emergence of In-Context Reinforcement Learning.*
    Jiuqi Wang, Rohan Chandra, **Shangtong Zhang**.
    Conference on Neural Information Processing Systems (**NeurIPS**), 2025
    Acceptance rate: 24.52%

28. *Finite Sample Analysis of Linear Temporal Difference Learning with Arbitrary Features.*
    Zixuan Xie*, Xinyu Liu*, Rohan Chandra, **Shangtong Zhang**.
    Conference on Neural Information Processing Systems (**NeurIPS**), 2025
    Acceptance rate: 24.52%

27. *Counterfactual Explanations for Continuous Action Reinforcement Learning.*
    Shuyang Dong, **Shangtong Zhang**, Lu Feng.
    International Joint Conference on Artificial Intelligence (**IJCAI**), 2025
    Acceptance rate: 19.3%

26. *Towards Large Language Models that Benefit for All: Benchmarking Group Fairness in Reward Models.*
    Kefan Song, Jin Yao, Runnan Jiang, Rohan Chandra, **Shangtong Zhang**.
    Reinforcement Learning Conference (**RLC**), 2025
    Acceptance rate: 39%

25. *Linear Q-Learning Does Not Diverge in $L^2$: Convergence Rates to a Bounded Set.*
    Xinyu Liu*, Zixuan Xie*, **Shangtong Zhang**.
    International Conference on Machine Learning (**ICML**), 2025.
    Acceptance rate: 26.9%

24. *Transformers Can Learn Temporal Difference Methods for In-Context Reinforcement Learning.*
    Jiuqi Wang*, Ethan Blaser*, Hadi Daneshmand, **Shangtong Zhang**.
    International Conference on Learning Representations (**ICLR**), 2025.
    Acceptance rate: 32.08%.
    **QuantCo Spotlight Award** at the ICML Workshop on In-Context Learning, 2024.

23. *Revisiting a Design Choice in Gradient Temporal Difference Learning.*
    Xiaochi Qian, **Shangtong Zhang**.
    International Conference on Learning Representations (**ICLR**), 2025.
    Acceptance rate: 32.08%.

22. *Efficient Policy Evaluation with Safety Constraint for Reinforcement Learning.*
    Claire Chen*, Shuze Liu*, **Shangtong Zhang**.
    International Conference on Learning Representations (**ICLR**), 2025.
    Acceptance rate: 32.08%.

21. *Doubly Optimal Policy Evaluation for Reinforcement Learning.*
    Shuze Liu, Claire Chen, **Shangtong Zhang**.
    International Conference on Learning Representations (**ICLR**), 2025.
    Acceptance rate: 32.08%.

20. *Efficient Multi-Policy Evaluation for Reinforcement Learning.*
    Shuze Liu, Claire Chen, **Shangtong Zhang**.

AAAI Conference on Artificial Intelligence (**AAAI**), 2025.
Acceptance rate: 23.4%. **Oral presentation (4.6%)**.

19. *Efficient Policy Evaluation with Offline Data Informed Behavior Policy Design.*
<u>Shuze Liu</u>, **Shangtong Zhang**.
International Conference on Machine Learning (**ICML**), 2024.
Acceptance rate: 27.5%

18. *On the Convergence of SARSA with Linear Function Approximation.*
**Shangtong Zhang**, Remi Tachet des Combes, Romain Laroche.
International Conference on Machine Learning (**ICML**), 2023.
Acceptance rate: 28%

17. *A Deeper Look at Discounting Mismatch in Actor-Critic Algorithms.*
**Shangtong Zhang**, Romain Laroche, Harm van Seijen, Shimon Whiteson, Remi Tachet des Combes.
International Conference on Autonomous Agents and Multiagent Systems (**AAMAS**), 2022.
Acceptance rate: 26%. **Oral presentation.**

16. *Learning Expected Emphatic Traces for Deep RL.*
Ray Jiang, **Shangtong Zhang**, Veronica Chelu, Adam White, Hado van Hasselt.
AAAI Conference on Artificial Intelligence (**AAAI**), 2022.
Acceptance rate: 15%.

15. *Breaking the Deadly Triad with a Target Network.*
**Shangtong Zhang**, Hengshuai Yao, Shimon Whiteson.
International Conference on Machine Learning (**ICML**), 2021.
Acceptance rate: 21.5%.

14. *Average-Reward Off-Policy Policy Evaluation with Function Approximation.*
**Shangtong Zhang**$^*$, Yi Wan$^*$, Richard S. Sutton, Shimon Whiteson.
International Conference on Machine Learning (**ICML**), 2021.
Acceptance rate: 21.5%.

13. *Mean-Variance Policy Iteration for Risk-Averse Reinforcement Learning.*
**Shangtong Zhang**, Bo Liu, Shimon Whiteson.
AAAI Conference on Artificial Intelligence (**AAAI**), 2021.
Acceptance rate: 21.4%.

12. *Learning Retrospective Knowledge with Reverse Reinforcement Learning.*
**Shangtong Zhang**, Vivek Veeriah, Shimon Whiteson.
Conference on Neural Information Processing Systems (**NeurIPS**), 2020.
Acceptance rate: 20.1%.

11. *GradientDICE: Rethinking Generalized Offline Estimation of Stationary Values.*
**Shangtong Zhang**, Bo Liu, Shimon Whiteson.
International Conference on Machine Learning (**ICML**), 2020.
Acceptance rate: 21.8%.

10. *Provably Convergent Two-Timescale Off-Policy Actor-Critic with Function Approximation.*
**Shangtong Zhang**, Bo Liu, Hengshuai Yao, Shimon Whiteson.
International Conference on Machine Learning (**ICML**), 2020.
Acceptance rate: 21.8%.

9. *Deep Residual Reinforcement Learning.*
**Shangtong Zhang**, Wendelin Boehmer, Shimon Whiteson.
International Conference on Autonomous Agents and Multiagent Systems (**AAMAS**), 2020.
Acceptance rate: 23%. **Best Paper Award**.

8. *Mega-Reward: Achieving Human-Level Play without Extrinsic Rewards.*
Yuhang Song, Jianyi Wang, Thomas Lukasiewicz, Zhenghua Xu, **Shangtong Zhang**, Andrzej Wojcicki, Mai Xu.

AAAI Conference on Artificial Intelligence (**AAAI**), 2020.
Acceptance rate: 20.6%.

7. *DAC: The Double Actor-Critic Architecture for Learning Options.*
**Shangtong Zhang**, Shimon Whiteson.
Conference on Neural Information Processing Systems (**NeurIPS**), 2019.
Acceptance rate: 21.2%.

6. *Generalized Off-Policy Actor-Critic.*
**Shangtong Zhang**, Wendelin Boehmer, Shimon Whiteson.
Conference on Neural Information Processing Systems (**NeurIPS**), 2019.
Acceptance rate: 21.2%.

5. *Distributional Reinforcement Learning for Efficient Exploration.*
Borislav Mavrin, **Shangtong Zhang**, Hengshuai Yao, Linglong Kong, Kaiwen Wu, Yaoliang Yu
International Conference on Machine Learning (**ICML**), 2019.
Acceptance rate: 22.6%.

4. *ACE: An Actor Ensemble Algorithm for Continuous Control with Tree Search.*
**Shangtong Zhang**, Hao Chen, Hengshuai Yao.
AAAI Conference on Artificial Intelligence (**AAAI**), 2019.
Acceptance rate: 16.2%. **Spotlight presentation.**

3. *QUOTA: The Quantile Option Architecture for Reinforcement Learning.*
**Shangtong Zhang**, Borislav Mavrin, Linglong Kong, Bo Liu, Hengshuai Yao.
AAAI Conference on Artificial Intelligence (**AAAI**), 2019.
Acceptance rate: 16.2%. **Oral presentation.**

2. *Crossprop: Learning Representations by Stochastic Meta-Gradient Descent in Neural Networks.*
Vivek Veeriah*, **Shangtong Zhang***, Richard S. Sutton.
European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in
Databases (**ECML-PKDD**), 2017.
Acceptance rate: 27.1%.

1. *A Deep Neural Network for Modeling Music.*
Pengjing Zhang, Xiaoqing Zheng, Wenqiang Zhang, Siyan Li, Sheng Qian,
Wenqi He, **Shangtong Zhang**, Ziyuan Wang
International Conference on Multimedia Retrieval (**ICMR**), 2015.
Acceptance rate: 31%.

**Refereed Workshop Papers (Non-Archival)**

4. *CRASH: Challenging Reinforcement-Learning Based Adversarial Scenarios For Safety Hardening.*
Amar Kulkarni, **Shangtong Zhang**, Madhur Behl.
ML for Autonomous Driving Workshop at AAAI, 2025.

3. *A Deeper Look at Experience Replay.*
**Shangtong Zhang**, Richard S. Sutton.
Deep RL Symposium at NIPS, 2017.

2. *Comparing Deep Reinforcement Learning and Evolutionary Methods in Continuous Control.*
**Shangtong Zhang**, Osmar R. Zaiane
Deep RL Symposium at NIPS, 2017

1. *A Demon Control Architecture with Off-Policy Learning and Flexible Behavior Policy.*
**Shangtong Zhang**, Richard S. Sutton.
Hierarchical RL Workshop at NIPS, 2017.

**Technical Reports**

1. *Group Fairness in Multi-Task Reinforcement Learning.*
   Kefan Song, Runnan Jiang, Rohan Chandra, **Shangtong Zhang**.
   ArXiv Preprint, 2025.


# GRANTS

9. Towards Safe and Efficient Automated Vehicle Platooning through
   Reinforcement Learning from Virtual Reality.
   CCI-CVN-UVA, **PI**, Total $50,000, My Share $25,000                    2026

8. Continual Skill Acquisition through In-Context Reinforcement Learning.
   Cisco Research, **PI**, Total $75,000, My Share $75,000                 2025 - 2026

7. Reinforced Sensory Intelligence for Safe Autonomous Driving.
   CCI Central Virginia Node, **PI**, Total $100,000, My Share $60,000     2025 - 2026

6. CAREER: Revolutionizing the Evaluation of AI Agents with Online and Offline Data.
   NSF 2442098, **PI**, Total $600,000, My Share $600,000                  2025 - 2030

5. Revolutionizing Domain Generalization with In-Context Reinforcement Learning.
   Nvidia Academic Grant, **PI**, 16,000 A100 Hours ($\sim$$64,000)        2025

4. Enhancing the Security of Large Language Models Against Persuasion-Based Jailbreak Attacks
   in Multi-Turn Dialogues.
   CCI Coastal Virginia Node, **Co-PI**, Total $60,000, My Share $20,000   2025 - 2026

3. RAMPART: Reinforcement Against Malicious Penetration by Adversaries in Realistic Topologies.
   DARPA HR001123S0002, **Co-PI** of UVA subaward, Total $2,150,000, My Share $77,000  2023 - 2027

2. SLES: CRASH: Challenging Reinforcement-Learning Based Adversarial Scenarios
   for Safety Hardening.
   NSF 2331904, **Co-PI**, Total $800,000, My Share $400,000               2023 - 2026

1. III: Small: Moving Offline Learning to Rank Online, from Theory to Practice.
   NSF 2128019, **PI**, Total $500,000, My Share $500,000                  2021 - 2026


# SERVICES

**Organizers**
CPS Rising Star Workshop 2024, Co-Chair

**Grant Panelist and Reviewer**
Austrian Science Fund (FWF), 2025
National Science Foundation (NSF), 2024
Virginia's Commonwealth Cyber Initiative (CCI), 2024

**Meta Meta Reviewer**
RL Conference 2025, 2026 (Senior Area Chair)

**Meta Reviewer**
NeurIPS 2025 (Area Chair)
ICML 2025, 2026 (Area Chair)
ICLR 2024, 2025, 2026 (Area Chair)
AISTATS 2024, 2025, 2026 (Area Chair)
L4DC 2025 (Program Committee)
AAMAS 2025 (Senior Program Committee)

RL Conference 2024 (Area Chair)
ACML 2022, 2023, 2024 (Area Chair)

**Reviewer**
Expert Reviewer for TMLR Outstanding Certification 2025
Transactions on Pattern Analysis and Machine Intelligence (1)
Transaction of Machine Learning Research (2)
Journal of Machine Learning Research (3)
Artificial Intelligence Journal (2)
Transactions on Intelligent Systems and Technology (2)
IJCAI 2023
AISTATS 2022
NeurIPS 2020, 2021, 2022, 2023
ICML 2020, 2021, 2022, 2023
AAAI 2020, 2021, 2022, 2023
ICLR 2021, 2022, 2023
SIGCOMM 2022
Offline Reinforcement Learning Workshop at NeurIPS 2020, 2021, 2022
Deep Reinforcement Learning Workshop at NeurIPS 2019, 2020, 2021, 2022
Adaptive and Learning Agents Workshop at AAMAS 2019, 2020
Optimization Foundations for Reinforcement Learning Workshop at NeurIPS 2019
Reinforcement Learning for Real Life Workshop at ICML 2019, 2021
Reinforcement Learning for Real Life Workshop at NeurIPS 2022

**Conference Session Chair**
AAAI 2023, "Reinforcement Learning Theory & Algorithms"


# SUPERVISION

**Doctral Students**

| | |
|---|---|
| Ethan Blaser, NSF Graduate Research Fellowship | 2023 - Now |
| Jiuqi Wang | 2023 - Now |
| Xinyu Liu | 2024 - Now |
| Zixuan Xie, UVA Provost's Fellowship, UVA Engineering Distinguished Fellowship | 2024 - Now |
| Amir Moeini | 2024 - Now |

**Alumni with Theses or Publications**

| | |
|---|---|
| Shuze Liu, PhD → Research Scientist at Meta | 2022 - 2025 |
| Kefan Song, Master → PhD at UVA | 2023 - 2025 |
| Licheng Luo, Master → PhD at UC Riverside | 2023 - 2024 |
| Claire Chen, Undergraduate → Undergraduate at Caltech | 2023 - 2025 |
| Vikram Ostrander, Undergraduate (DMP thesis) → Software Engineer at Palantir | 2024 - 2025 |
| Vagul Mahadevan, Undergraduate (DMP thesis) → Software Engineer at Metron | 2023 - 2025 |
| Steve Zhou, Undergraduate (DMP thesis) → Master at MIT | 2023 - 2024 |
| Xiaochi Qian, Undergraduate (Oxford) → Quantitative Researcher at Marshall Wace | 2022 - 2024 |

**PhD Dissertation Examination Committees**
Shuyang Dong (by Prof. Lu Feng)
Qianhong Zhao (by Prof. Gang Tao)
Fengdi Che (U of Alberta, by Prof. Rupam Mahmood and Prof. Dale Schuurmans)
Shohaib Mahmud (by Prof. Haiying Shen)
Suraiya Tairin (by Prof. Haiying Shen)
Tao Jin (by Prof. Farzad Farnoud)

Fan Yao (by Prof. Hongning Wang and Prof. Haifeng Xu)
Ingy ElSayed-Aly (by Prof. Lu Feng)
Sudhir Shenoy (by Prof. Afsaneh Doryab)
Chuanhao Li (by Prof. Hongning Wang)

## INVITED TALKS

**Convergence of Value-Based Reinforcement Learning: Advances and Open Problems**
Workshop on Advances in ML Theory, Canada, by Prof. Csaba Szepesvari — May 2025
**Understanding the Training and Inference of Reinforcement Learning**
Tsinghua University, hosted by Prof. Hongning Wang — June 2024
**On the Cheating of Offline Reinforcement Learning**
KAUST Rising Stars in AI Symposium — Feb 2024
**Offline Reinforcement Learning: Current and Future**
AAAI New Faculty Highlight Program — Feb 2023
**Breaking the Deadly Triad in Off-Policy Reinforcement Learning**
Department of Computer Science, University of Virginia — Mar 2022
School of Computing Science, Simon Fraser University — Feb 2022
Department of Electrical & Computer Engineering, University of Waterloo — Feb 2022
School of Informatics, University of Edinburgh — Oct 2021
**Breaking the Deadly Triad with a Target Network**
Microsoft Research Summit — Oct 2021
**Breaking the Deadly Triad in Reinforcement Learning**
RL team, DeepMind, hosted by Dr. Hado van Hasselt — Sep 2021
**Off-Policy Evaluation**
Data Fest 2020, Open Data Science — Oct 2020
**Off-Policy Evaluation and Control**
ByteDance AI Lab, Shanghai — Oct 2020
**Coding Deep RL Papers**
NIPS MLTrain Workshop, Long Beach — Dec 2019
**Off-Policy Actor-Critic Algorithms**
Latent Logic LTD, Oxford — Apr 2019

## TEACHING & OUTREACH

**University of Virginia**
CS 6771: Reinforcement Learning — Spring 2026
CS 4771: Reinforcement Learning — Fall 2025
CS 6316: Machine Learning — Spring 2024, Spring 2025
CS 4501: Reinforcement Learning — Fall 2024
CS 4501: Optimization — Fall 2023
CS 6501: Topics in Reinforcement Learning — Fall 2022
**Tutorial: What GenAI Can and Cannot Do?**
UVA Innovation Hub at Charlottesville Middle School — July 2025
**Charlottesville High School – Link Lab Mentorship Program**
Gravity Battery & Wearable GPS — Fall 2024

## OPEN SOURCE CODE

**GitHub Repo: PyTorch Deep RL**
A zoo of popular deep RL algorithms in PyTorch with **3k stars**.

**GitHub Repo: Reinforcement Learning: An Introduction**
Python implementation of the book *Reinforcement Learning: An Introduction* with **13.8k stars**.

**Google Summer of Code (GSoC)**

| | |
|---|---|
| MLPack | 2017 |
| The Xapian Project | 2014 |