

Alex Ayoub

1–36 Athabasca Hall, University of Alberta – Edmonton, AB – Canada

☎ +1 561 210 3181 • ✉ aayoub@ualberta.ca • Citizenship: United States

Education

University of Alberta

Ph.D. in Computing Science

Supervisors: Csaba Szepesvári & Dale Schuurmans

Edmonton, AB, Canada

Sep 2021–May 2026 (Expected)

University of Alberta

M.Sc. in Computing Science

Supervisor: Csaba Szepesvári

Thesis: Towards Sample Efficient Reinforcement Learning with Function Approximation

Nominated for Outstanding Thesis Award

Edmonton, AB, Canada

Sep 2019–Sep 2021

Florida State University

B.Sc. in Computational Science and Applied Mathematics

Tallahassee, FL, USA

Jun 2015–May 2019

Preprints

- Pretraining large language models with reinforcement learning to overcome the inherent limitations of canonical next token prediction. Current results indicate that our method overcomes failure cases of next token prediction, such as star graphs, while maintaining and enhancing performance on language tasks such as the Tiny Stories dataset by enabling the model to either complete or begin stories from a prompt. *In preparation*

Publications

- **A. Ayoub**, K. Asadi, D. Schuurmans, C. Szepesvári, K. Bouyarmane (2026). *Learning to Reason Efficiently with Discounted Reinforcement Learning*. **ICLR**.
- A. Bakhtiari*, **A. Ayoub***, S. Robertson, D. Janz, C. Szepesvári (2025). *Eluder dimension: localise it!* **NeurIPS** [Spotlight; top 3%].
- **A. Ayoub***, D. Szepesvári*, A. Bakhtiari, C. Szepesvári, D. Schuurmans (2025). *Rectifying Regression in Reinforcement Learning*. **Reinforcement Learning Conference (RLC)**.
- **A. Ayoub**, S. Robertson, D. Liang, H. Steck, N. Kallus (2025). *Does weighting improve matrix factorization for recommender systems?* **The Web Conference (WWW)**.
- S. Liu*, **A. Ayoub***, F. Sentenac, X. Tan, C. Szepesvári (2024). *Almost Free: Self-concordance in Natural Exponential Families and an Application to Bandits*. **NeurIPS**.
- **A. Ayoub**, D. Szepesvári, F. Zanini, B. Chan, D. Gupta, B. Castro da Silva, D. Schuurmans (2024). *Mitigating the Curse of Horizon in Monte-Carlo Returns*. **RLC**.
- **A. Ayoub**, K. Wang, V. Liu, S. Robertson, J. McInerney, D. Liang, N. Kallus, C. Szepesvári (2024). *Switching the Loss Reduces the Cost in Batch Reinforcement Learning*. **ICML**.
- D. Janz*, S. Liu*, **A. Ayoub***, C. Szepesvári (2024). *Exploration via linearly perturbed loss minimisation*. **International Conference on Artificial Intelligence and Statistics (AISTATS)** [Oral; top 1%].
- Z. Zhang, J. Kirschner, J. Zhang, F. Zanini, **A. Ayoub**, D. Schuurmans (2023). *Managing Temporal Resolution in Continuous Value Estimation: A Fundamental Trade-off*. **NeurIPS**.
- E. Miah, R. MacQueen, **A. Ayoub**, A. Masoumzadeh, M. White (2023). *Resmax: An Alternative Soft-Greedy Operator for Reinforcement Learning*. **Transactions on Machine Learning Research (TMLR)**.
- H. Ishfaq*, Q. Cui*, V. Nguyen*, **A. Ayoub***, Z. Yang, Z. Wang, D. Precup, L. F. Yang (2021). *Randomized Exploration for Reinforcement Learning with General Value Function Approximation*. **ICML**.
- **A. Ayoub**, Z. Jia, C. Szepesvári, M. Wang, L. Yang (2020). *Model-Based Reinforcement Learning with Value-Targeted Regression*. **ICML**.

* Equal contribution.

Professional Experience

Google DeepMind

Student Researcher — Frontier AI Unit

London, UK
Aug 2025–Present

- Pretraining large language models with reinforcement learning.

Amazon

Applied Scientist Intern — Generative AI & LLM Reasoning

Seattle, WA, USA
Jun 2025–Aug 2025

- Post-training large language models to reason more efficiently in theory and practice; work accepted at ICLR.

Netflix

Machine Learning Research Intern

Los Gatos, CA, USA
Jun 2024–Nov 2024

- Identified a major flaw in a critical in-production recommender component and proposed a simple, effective fix.
- Showed that a commonly used weighting practice can be removed without harming performance, simplifying deployment. Work accepted at WWW
- Designed a novel loss for decision-making that reduces data needs relative to standard losses.

Morgan Stanley

Machine Learning Research Intern

New York, NY (remote)
Feb 2024–May 2024

- High-frequency trading and market estimation in limit order books with RL.

Netflix

Machine Learning Research Intern

Los Gatos, CA, USA
Jun 2023–Aug 2023

- Demonstrated theoretical and empirical benefits of using log-loss in offline RL and recommender systems. Worked accepted at ICML

Spotify

Research Scientist Intern

New York, NY, USA
Jun 2022–Aug 2022

- Developed an online RL agent to calibrate Spotify homepage content, outperforming offline baselines in production.

Huawei Noah's Ark Lab

Research Scientist Intern

Edmonton, AB, Canada
Oct 2021–May 2023

University of Alberta

Teaching Assistant

Edmonton, AB, Canada
Sep 2019–Present

Technical Skills

Programming: Python, C/C++, MATLAB **Frameworks & Tools:** PyTorch, JAX, CuPy, Mathematica

Service, Awards, and Honors

- Inaugural recipient of the **Netflix Graduate Research Fellowship**.
- **Co-organizer**, RL Theory Workshop (2023, 2024).
- **Workflow Chair**, ICML 2022.
- FSU President's List (2018).
- University of Alberta Graduate Research Assistant Fellowship.
- Florida State University Research Assistant Grant.