

Mehrdad Moghimi

✉ mehrdad.m7496@gmail.com | 🌐 mehrdadmoghimi.github.io
🌐 mehrdad-moghimi | 👤 MehrdadMoghimi | 🎓 Google Scholar

SUMMARY

I am broadly interested in reinforcement learning as a framework for building agents that make reliable decisions under uncertainty and over long horizons. My early work focused on the theory of risk-sensitive RL, where I studied time inconsistency and developed principled algorithms to address it (ICML 2025). I then extended these ideas to larger-scale problems through actor-critic methods for both online and offline learning, implemented in JAX and PyTorch. Recently, I developed a constrained RL framework for safety-critical decision making, enabling safety budgets to be specified at test time and achieving state-of-the-art performance. Currently, I am extending my RL background to large language models, focusing on reinforcement learning for reasoning and knowledge distillation.

EDUCATION

Ph.D. in Applied Mathematics, York University, Toronto, Canada 09/2021 – 12/2026 (Expected)

- Supervisor: Prof. Hyejin Ku
- Research focus: Safe and Risk-sensitive Reinforcement Learning, Machine Learning, Quantitative Finance

MBA in Finance, Sharif University of Technology, Tehran, Iran 09/2018 – 09/2021

- GPA: 18.62/20
- Supervisor: Prof. Hamid Arian
- Research focus: Machine Learning, Quantitative Finance

BSc in Computer Science, Sharif University of Technology, Tehran, Iran 09/2014 – 09/2018

- GPA: 18.71/20 (Ranked 1st in Class)

ACADEMIC RESEARCH EXPERIENCE

Research Collaborator, with Bernardo Ávila Pires (Google DeepMind) 09/2025 – Present

- Utility-Constrained Policy Optimization | *Under review*
 - * Developed a practical framework for utility-constrained policy optimization using stock-augmented distributional RL.
 - * Designed a safe Lagrangian actor-critic algorithm with test-time controlled safety budgets.
 - * Implemented and evaluated the method on 22 Safety Gymnasium benchmarks, achieving state-of-the-art performance.

Research Assistant, York University, Toronto, ON, Canada 09/2021 – Present

- Theoretical Foundations of Risk-Sensitive Reinforcement Learning with Spectral Risk Measures | *ICML 2025*
 - * Studied the time-inconsistency phenomena in risk-sensitive reinforcement learning and developed a novel theory to characterize the behavior of risk-sensitive policies.
 - * Designed a time-consistent risk-sensitive RL algorithm and established convergence guarantees.
 - * Validated theoretical findings through empirical evaluation on various RL environments.
- Scalable Risk-Sensitive Actor-Critic for Online and Offline RL | *Expert Systems with Applications*
 - * Expanded the risk-sensitive framework with Spectral risk measures to high-dimensional problems by formulating Actor-Critic architectures suitable for both online and offline learning.
 - * Established convergence guarantees for the policy gradient updates in the tabular setting.
 - * Extended a risk-neutral JAX codebase to incorporate risk-sensitive objectives, and conducted extensive empirical evaluations demonstrating the algorithm's performance in both online and offline settings.
- Joint Analysis of Time and Risk Preferences in Reinforcement Learning | *Under Review*
 - * Investigated the interplay between agent time-preferences (via general discount functions) and risk-sensitivity to create more human-aligned agents.
 - * Demonstrated that existing frameworks for hyperbolic discounting can induce time-inconsistent behavior, leading to sub-optimal policies.

- * Developed a time-consistent formulation, resulting in significant performance improvements over standard hyperbolic baselines in various environments, including Atari.

Research Assistant, RiskLab, Tehran, Iran 09/2019 – 08/2021

- Developed “Encoded Value-at-Risk,” a novel machine learning framework utilizing Variational Auto-Encoders (VAEs) to model non-linear dependencies in high-dimensional financial data.

Research Assistant, Sharif Image Processing Lab, Tehran, Iran 01/2018 – 08/2018

- Bachelor’s Thesis: Designed a computer vision pipeline for dynamic advertisement replacement in broadcast soccer footage under Prof. Shohreh Kasaei.
- Implemented homography estimation to localize field banners, enabling realistic content substitution that accounts for camera perspective.

PUBLICATIONS

- **M. Moghimi**, B. A. Pires, “Utility-Constrained Policy Optimization”, Under review
- **M. Moghimi**, A. Coache, H. Ku, “Decoupling Time and Risk: Risk-Sensitive Reinforcement Learning with General Discounting”, Under review ([arXiv Link](#))
- **M. Moghimi**, H. Ku, “Risk-sensitive Actor-Critic with Static Spectral Risk Measures for Online and Offline Reinforcement Learning”, Published at at *Expert Systems with Applications* ([Link](#))
- **M. Moghimi**, H. Ku, “Beyond CVaR: Leveraging Static Spectral Risk Measures for Enhanced Decision-Making in Distributional Reinforcement Learning”, Published at *ICML 2025* ([Link](#))
- H. Arian, **M. Moghimi**, E. Tabatabaei, and S. Zamani, “Encoded Value-at-Risk: A machine learning approach for portfolio risk measurement”, Published at *Mathematics and Computers in Simulation, 2022* ([Link](#))

PROFESSIONAL EXPERIENCE

Machine Learning Research Intern, RBC Borealis, Toronto, Canada 05/2026 – Present

- Conducting research on knowledge distillation for large language models.
- Developing and evaluating novel distillation algorithms for efficient model training and inference.

Model Validation Intern, Sun Life Financial, Toronto, Canada 09/2022 – 12/2022

- Audited quantitative risk models, ensuring alignment with regulatory standards.
- Performed stress-testing to assess model robustness under extreme scenarios.
- Conducted numerical verification and sensitivity analysis on actuarial valuations, identifying discrepancies in risk metric calculations.

HONOURS AND AWARDS

- **York University Graduate Scholarship**, Fall 2021
- **Exceptional Talents Scholarship**, Direct admission to graduate studies (waived National Entrance Exam) due to top academic performance.

TECHNICAL SKILLS

- **Languages (Proficient):** Python
- **Languages (Intermediate):** R, MATLAB, Java
- **Frameworks:** PyTorch, JAX (Flax, Optax), Pandas