

# Erfan Miahi — RL Research Engineer

Canada – Toronto

✉ mhi.erfan1@gmail.com

☎ (+1) 8254360215

👤 [erfanmiahi.com](https://erfanmiahi.com)

## HIGHLIGHT OF SKILLS & QUALIFICATIONS

- 9+ years in AI research: Published 8 top-tier papers with ~200 citations, including EMNLP, AIJ, MLJ, and TMLR.
- 5+ years in MLE/SWE: Built scalable Reinforcement Learning (RL) pipelines for impactful real-world applications.
- Collaborations with top ML labs: M.Sc. at Richard Sutton's RLAI Lab; collaborated with Google DeepMind.
- End-to-end understanding of training models: Built GPU kernels in Triton and distributed RL training systems.
- Exceptional communication skills: Excel at mentoring diverse individuals across technical and business domains.

## KEY TECHNICAL SKILLS

Languages	Python (Advanced), C++, Triton, SQL, JavaScript, TypeScript, Java
ML Frameworks	PyTorch, TensorFlow, Jax, TRL, Transformers, PEFT (LoRA), vLLM, Unsloth (Kernels)
Distributed Training / HPC	DDP, FSDP, DeepSpeed, Megatron-LM, Accelerate, trlX, verifiers, Ray (RLlib, Train)
Infrastructure	Kubernetes, Docker, Slurm, Kubeflow, Ray Cluster, Terraform, PostgreSQL, Redis, S3
Experimentation & MLOps	W&B, GitHub Actions, Hydra, Pytest, Optuna, MLflow, Langfuse
Cloud	GCP (GKE, CE), AWS (EKS, ECS, EC2), Azure, Compute Canada HPC, Nebius
Foundations	ML/RL/DL, LLMs/VLMs, RLHF/RLVR, RAG, Math, Algorithms & Data Structures
Methodologies	Agile, MLOps, CI/CD, TDD, OOP, Design Patterns

## RELEVANT EXPERIENCES

**Technical Co-Founder (Currently Lead Research Engineer, CovenantAI)** **Aug 2023 – Present (2 yrs)**  
*Qelix (Website)* *Toronto, Canada*

- Co-founded a startup focusing on post-training LLMs for specialized coding/math/gpu-kernel agents using RL and SFT.
- Led R&D on off-policy RLVR, stabilizing 16-step policy staleness and enabling fully async inference-trainer updates.
  - Collaborated on pre-training a 72B LLM, the largest ever on decentralized infra ([Model](#), [Report](#), [1M+ views](#)).
  - Proposed a new inference verification algorithm, shrinking verification payloads to 4 bytes per token (2× gain).
  - Pioneered an RLVR/RLHF-aware weight communication protocol with ~100× bandwidth reduction.
  - Built q-code eval harness from scratch and post-trained LLMs on it; +40 pp pass@1 accuracy ([Blog](#)).
  - Led development of a decentralized distributed post-training system, achieving 2% → ~50% gain on MATH ([Blog](#)).
  - Accepted to Antler & NextAI (<3-5% admit); interviewed by Y Combinator (<7% admit).

**Machine Learning Research Engineer** **Nov 2022 – Oct 2024 (2 yrs)**  
*DeepR Analytics (Proprietary Trading Firm)* *Remote, Canada & US*

- Founding ML research engineer conducted research and built large-scale auto-trading pipeline and research tooling.
- Built PyTorch pipeline that trained >1 M RL agents, adding \$35+ daily PnL/asset/share.
  - Implemented SOTA papers & proposed RL algorithms in 10+ research projects; added \$20+ daily PnL/asset/share.
  - Migrated training/backtest stack to AWS ECS & Fargate, enabling on-demand scaling across 100+ assets.
  - Optimized data pipeline, simulated environments, eval pipeline, async I/O; cut epoch time 9x and EC2 cost 4x.
  - Raised test coverage to 85% with Pytest; slashed prod bugs 45% across ECS, CI, and C++ engine.

**Research Associate** **Sep 2022 - Dec 2022 (4 months)**  
*University of Alberta (RLAI Lab)* *Edmonton, Canada*

- Redesigned RL codebase for sensory data prediction, accelerating experimentation and analysis by 50% ([Publication](#)).
- Developed and deployed optimized offline RL and control algorithms, improving system performance by 21%.

**Assistant AI Consultant** **Nov 2021 - Mar 2023 (1 year 5 months)**  
*Alberta Machine Intelligence Institute (Amii)* *Edmonton, Canada*

- Provided continuous technical coaching to 5 enterprise teams, aligning projects with their business needs.
- Consulted an energy firm on problem formulation and ML design, potentially saving \$10-15M annually ([Blog](#)).

**AI Mentor (For More Info)** **Jan 2017 - present (7+ years)**  
*Self-employed (Leadership)* *Remote/Edmonton, Canada*

- Mentoring 10+ students (past and present) in AI research and engineering for 7+ years.
- Mentored a University of Guilan (rank 1422) student to secure a PhD at Michigan State University (rank 116).

**Software Developer** **Nov 2015 – Nov 2017 (2+ years)**  
*Nilasoft* *Rasht, Iran*

- Built and optimized a Spring/Angular e-commerce platform for 1K+ users with 99.9% uptime.

## EDUCATION

---

**M.Sc. in Computing Science**

*University of Alberta*

**Jan 2021 - Aug 2022**

*Edmonton, Canada*

○ **Supervisors:** *Martha White & Marlos C. Machado*

○ **Thesis:** Feature Generalization in Deep RL: An Investigation into Representation Properties

○ **Committee:** *Martha White (UoA) & Marlos C. Machado (Ex DeepMind) & Matthew E. Taylor (Ex BorealisAI)*

**B.Sc in Computer Engineering**

*University of Guilan*

**Sep 2015 - Aug 2019**

*Rasht, Iran*

○ **Ranked 1<sup>st</sup> in class**

## RELEVANT PROJECTS (60+ GitHub Projects)

---

**Grail: Open-Source Distributed, Decentralized, and Incentivized Post-Training System**

[Repo](#)

**A Composable Reinforcement Learning Pipeline for Post-training LLMs for Reasoning**

[Repo](#)

**AI-first Software Development Copilot for Specialized Industries (e.g., Optics)**

[Backend](#)

**Distributed RL Framework: A3C/PPO/IMPALA Implementations**

[Repo](#)

**Open-Source RL Framework for Fast Experimentation on Real-World Problems (Sole Contributor)**

[Repo](#)

**Circuit Sketch to VR: Meta Quest 3 Circuit Analyzer (Meta Hackathon 1st Place)**

[Backend](#) / [Frontend](#)

**Mixed-Precision Library for LLM Training with Torch JIT and Bitwise Manipulation (FP8 and BF16)**

**Distributed LLM Training Infrastructure using Kubernetes and Terraform (DDP/FSDP in PyTorch)**

**Solutions to Assignments of Deep RL Course by UoC, Berkeley (100+ Stars on GitHub)**

[TensorFlow](#) / [PyTorch](#)

## SELECTED PUBLICATIONS (~200 Citations on Google Scholar)

---

**Covenant-72B: Pre-Training a 72B LLM with Trustless Peers Over-the-Internet**

Joel Lidin, Amir Sarfi, Erfan Miah, et al.

[Paper](#)

*arXiv Preprint (1M+ views on Twitter)*

**Calibrating LLM Confidence by Probing Perturbed Representation Stability**

Reza Khanmohammadi, Erfan Miah, et al.

[Paper](#)

*Published as a main conference paper at EMNLP (nominated for outstanding paper award)*

**Investigating the Properties of Neural Network Representations in Reinforcement Learning**

Han Wang, Erfan Miah, Martha White, Marlos C. Machado, et al. (DeepMind Collaboration)

[Paper](#)

*Published in Artificial Intelligence Journal (AIJ)*

**GVPs in the Real World: Making Predictions Online for Water Treatment**

Muhammad K. Janjua, Haseeb Shah, Martha White, Erfan Miah, et al.

[Paper](#)

*Published in Machine Learning Journal (MLJ)*

**Scalable Transfer Evolutionary Optimization: Coping with Big Task Instances**

Mojtaba Shakeri, Erfan Miah, Abhishek Gupta, Yew-Soon Ong (NTU & A\*STAR Collaboration)

[Paper](#) / [Code](#)

*Published in IEEE Transactions on Cybernetics Journal*

**Understanding and Exploiting Weight Update Sparsity for Communication-Efficient Distributed RL**

Erfan Miah, Eugene Belilovsky

[Paper](#)

*arXiv Preprint*

## AWARDS & ACHIEVEMENTS

---

○ **1st Place at Llama Meta Hackathon Toronto 2024 (45 teams and 150 attendees) - Won \$25K, Unitree Robot, and 4 Ray-Bans**

○ **Reached Y Combinator Summer 2024 interview stage (top 7% applicant)**

○ **Placed among the top 10% of reviewers in ICML 2022**

○ **Ranked 6th in World Robot Challenge 2018 (received travel grant to Japan)**

○ **Abrishamchian Reward (Awarded for graduating with the highest distinction among all students of the University)**

○ **Top Researcher in Computer Engineering Group**

○ **Honorable Mention in ACM/ICPC 2016**

○ **Reviewer at ICML and TMLR**

## LANGUAGES

---

**English - Fluent**   **Persian - Native**