

Erfan Miah — ML Engineer

Canada – Toronto

✉ mhi.erfan1@gmail.com

☎ (+1) 8254360215

👤 erfanmiah.com

HIGHLIGHT OF SKILLS & QUALIFICATIONS

- **3+ years in MLE/SWE:** Built scalable AI systems and ML pipelines, delivering impactful real-world solutions.
- **8+ years in AI Research:** Published **6 top-tier papers** with **100+ citations**, collaborating with **Google DeepMind**.
- **End-to-end understanding of LLMs:** Wrote GPU kernels, implemented DDP, and developed AI agents with RAG.
- **Exceptional communication skills:** Excel at mentoring diverse individuals across technical and business domains.
- **Proven project management:** Successfully managed multiple projects while maintaining quality within timelines.

RELEVANT EXPERIENCES

Machine Learning Research Engineer

DeepR Analytics

November 2022 - October 2024 (2 years)

Remote, Canada & US

Part of the founding ML team developing a Deep Reinforcement Learning (RL) system for automated trading, where I:

- **Developed RL pipeline and algorithms with PyTorch**, driving **\$50+ daily profit increase per asset**.
- **Migrated pipeline to Amazon ECS**, enabling **large-scale deployment and experimentation** on various trading assets.
- **Optimized ML Pipeline for 9x faster model training**, reducing resource usage significantly.
- **Achieved 85% test coverage with Pytest**, reducing bugs by **45%** across ECS, GitHub CI, and C++ Trading Engine.

CEO

IntractAI (Leadership and Business)

August 2023 - August 2024 (1 year)

Toronto, Canada

- **Led a team of 2 engineers/researchers**, driving the development of backend infrastructure.
- Built backend infrastructure for **continual LLM fine-tuning** on the codebase and internal documents.
- Developed a **RAG pipeline** and **web-scraping agent** to improve code generation by **20%**.
- Optimized LLM models for **4x faster inference** using quantization, VLLM library, and flash-attention.

Research Associate

University of Alberta (RLAI Lab)

Sep 2022 - Dec 2022 (4 months)

Edmonton, Canada

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

Assistant AI Consultant

Alberta Machine Intelligence Institute (Amii)

Nov 2021 - Mar 2023 (1 year 5 months)

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](#))

Self-employed (Leadership)

Jan 2017 - present (7+ years)

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

Nilasoft

Oct 2016 – Nov 2017 (1 year 2 months)

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 1st in class**

RELEVANT PROJECTS ([60+ GitHub Projects](#))

Llama Circuit Analysis: Converting Sketches to Interactive Circuits in Meta Quest 3

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

LLM-Powered Contract Review and Extraction for Revenue Accountants

[Repo](#)

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

Backend

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

Request

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

Repo

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

TensorFlow/PyTorch

KEY SKILLS

ML Frameworks	PyTorch, TensorFlow, VLLM, LangChain, XGBoost, Scikit-learn, Optuna
Programming Languages	Python, C++, Java, JavaScript, TypeScript, SQL, HTML, CSS
Web Development	Angular, React, NextJS, FastAPI, Kubernetes
Cloud Services	Google Cloud, Compute Canada, AWS Cloud, Azure
Operating Systems	Windows, GNU/Linux (Ubuntu & Fedora & RHEL), macOS
Extra Tools	Ray-core, PostgreSQL, Redis, Docker, GitHub Actions, WandB, Git, Vim, ROS
Strong Background in	ML and Deep Learning, RL, Mathematics, Data Structures and Algorithms

SELECTED PUBLICATIONS (100+ Citations on Google Scholar)

Investigating the Properties of Neural Network Representations in Reinforcement Learning

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

Published in Artificial Intelligence Journal (AIJ)

Paper

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

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

Published in Machine Learning Journal (MLJ)

Paper

Scalable Transfer Evolutionary Optimization: Coping with Big Task Instances

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

Published in IEEE Transactions on Cybernetics Journal

Paper/Code

Resmax: An Alternative Soft-Greedy Operator for Reinforcement Learning

Erfan Miah, Revan MacQueen, Alex Ayoub, Abbas Masoumzadeh, Martha White

Published in Transactions in Machine Learning Journal (TMLR)

Paper

Genetic Neural Architecture Search for Automatic Assessment of Human Sperm Images

Miah, Erfan, Seyed Abolghasem Mirroshandel, and Alexis Nasr

Published in Expert Systems and Applications

Paper

AWARDS AND ACHIEVEMENTS

- 1st Place at Llama Meta Hackathon Toronto 2024 (45 teams and 150 attendees)
- 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

LANGUAGES

English - Fluent Persian - Native French - Novice