

Raihan Seraj

✉ raihan.seraj@mail.mcgill.ca 📞 514-998-8234 🌐 <https://www.linkedin.com/in/raihan-seraj/>
🌐 [raihan-seraj.github.io](https://github.com/raihan-seraj) 🌐 Raihan-Seraj 🎓 google scholar 🏠 Toronto, ON Canada

EDUCATION

McGill University, Montreal, Canada Jan 2019 – Apr 2025
PhD, Electrical and Computer Engineering
Thesis: Dynamic modeling of mental workload and fatigue-aware decision referrals in human automation teams
Supervisors: Aditya Mahajan, Jerome Le Ny.

McGill University, Montreal, Canada Jan 2017 – Jan 2019
MEng(Thesis), Electrical and Computer Engineering
Thesis: Learning in the presence of partial observability and concept drifts.
Supervisor: Aditya Mahajan

Islamic University of Technology, Dhaka, Bangladesh Dec 2011 – Dec 2015
BSc, Electrical and Electronic Engineering
Supervisor: Rakibul Islam

Affiliation: *Mila* (Quebec AI Institute).

RESEARCH INTEREST

Foundation-scale time-series modeling for high-dimensional dynamical systems.
Multi-agent reinforcement learning and emergent decentralized intelligence.
Human-in-the-loop AI, shared autonomy, and adaptive control.
Large-scale deep learning, diffusion models, and generative frontier architectures.

SELECTED PUBLICATIONS (A full list is available on Google Scholar.)

1. **Raihan Seraj**, Aditya Mahajan, Jerome Le Ny. “Fatigue and task load dependent decision referrals for joint binary classification in human-automation teams” *IEEE Control Systems Letters*, 2025
2. **Raihan Seraj**, Aditya Mahajan, Jerome Le Ny. “Dynamic Estimation of Mental Workload and Operator Accuracy in Human Automation Teams” *IEEE Transactions on Human Machine Systems*, 2025
3. Qi Yan, **Raihan Seraj**, Jiawei He, Lili Meng, Tristan Sylvain. “Autocast ++: Enhancing world event prediction with zero-shot ranking based context retrieval”. *International Conference on Learning Representation (ICLR)*, 2024.
4. Anurag Koul, Shivakanth Sujit, Shaoru Chen, Ben Evans, Lili Wu, Byron Xu, Rajan Chari, Riashat Islam, **Raihan Seraj**, Yonathan Efroni, Lekan P Molu, Miroslav Dudík, John Langford, Alex Lamb. “PcLast: Discovering Plannable Continuous Latent States”. *International Conference on Machine Learning (ICML)*, 2024.
5. **Raihan Seraj**, Jivitesh Sharma, Ole-Christoffer Granmo. “Tsetlin Machine for solving contextual bandit problems”. *35th Conference on Neural Information Processing System (NeurIPS)*, 2022
6. **Raihan Seraj**, Jerome Le Ny, Aditya Mahajan. “Mean field approximation for large population beauty contest games”. *IEEE Conference on Decision and Control (CDC)* 2021.
7. Jayakumar Subramanian, Amit Sinha, **Raihan Seraj**, Aditya Mahajan. “Approximate information state for approximate planning and reinforcement learning in partially observed environments”. *Journal of Machine Learning Research (JMLR)*, 2020.
8. Riashat Islam, **Raihan Seraj**, Pierre-Luc Bacon, Doina Precup. “Entropy Regularization with Discounted Future State Distribution in Policy Gradient Methods”. *NeurIPS Workshop on Optimization Foundations for Reinforcement Learning*, 2019.

RESEARCH EXPERIENCE DURING GRADUATE STUDIES

Research Assistant

Jan 2017 – April 2025

McGill University and Mila (Quebec AI Institute), Montreal, Canada

- Conducted research for graduate studies on partially observable Markov decision processes.
- Conducted research for graduate studies on learning in the presence of concept drifts for Wi-Fi localization.
- Conducted research for graduate studies on human in the loop control designed experimental protocols and hypothesized dynamic estimation of mental workload.
- Conducted research for graduate studies on decision referrals in human automation teams.

Research Intern

Aug 2023 – Feb 2024

Borealis AI, Royal Bank of Canada, Montreal, Canada

Mentors: Tristan Sylvain, Layla El Asri, Greg Mori

- Designed AI model for event forecasting from large-scale text corpora via AutoCast++ using zero-shot ranking and transformer-based context retrieval.
- Developed entropy-based human feedback algorithms for contextual bandits, enhancing decision accuracy and model performance for real-time recommendation systems.

Mila Scientist in Residence Program

June 2021–Sept 2021

Valence Drug Discovery, Montreal, Canada

Mentor: Daniel Cohen

- Designed fibrous autoencoder architectures to learn latent molecular spaces that enable the generation of structurally diverse compounds while maintaining consistent bio-activity profiles.
- Formulated a reinforcement learning framework for scaffold hopping using latent molecular representations, collaborating with the R&D team to efficiently explore large chemical spaces and accelerate de novo drug discovery.

Mila Scientist in Residence Program

March 2020 – Oct 2020

Paladin AI, Montreal, Canada

Mentor: Mikhail Klassen

- Studied pilot performance evaluation through feature-based segmentation of flight simulator data.
- Analyzed time series data from aircraft simulators to segment pilot reactions during faults, leading to the development of a proficiency metric for pilot performance evaluation.

Research Scientist Intern, Mitacs

March 2018 – May 2019

Aerial Technologies, Montreal, Canada

Mentors: Negar Ghourchian, Michel Allegue

- Designed feature extraction algorithms for WiFi Channel State Information (CSI) data, which improved model generalization by addressing concept drift in real-time indoor localization systems.
- Analyzed time series data and trained deep learning models to classify fall detection using WiFi Channel State Information (CSI).

PROFESSIONAL EXPERIENCE

Member of Technical Staff

Feb 2026 – Present

Bagel Labs Canada Inc, Toronto, Canada

- Implemented a frontier-scale video generation model using a decentralized diffusion framework, enabling training and inference without centralized data or parameter sharing.
- Designed and implemented text-to-image diffusion models trained entirely via decentralized computation, demonstrating scalable generative modeling under fully distributed, communication-free settings.
- Designed and implemented distributed reinforcement learning algorithm for fine-tuning large language models (LLMs) across globally distributed compute nodes, enabling scalable, high-throughput training and reducing convergence time by leveraging worldwide parallelism.
- Designed and built a robust reinforcement learning library for fine-tuning large language models (LLMs) on tool-use tasks.

Member of Technical Staff

May 2025 – Jan 2026

Schrodinger Journey Labs (Parent Company of Bagel Labs) (Remote), Montreal, Canada

- Conducted research in machine learning, including designing and running experiments, validating hypotheses, and stayed current with academic literature and industry developments.
- Trained, fine-tuned, and evaluated machine learning models and related systems
- Prepared datasets, including data collection, cleaning, preprocessing, labeling, augmentation, and quality control.

Senior ML Engineer (Part Time Contract)

Oct 2024 – Feb 2025

Braxted Inc (Remote), Montreal, Canada

- Worked with a team of engineers and researchers in the successful design, development, and deployment of agentic Retrieval-Augmented Generation (RAG) systems tailored for Canadian immigration processes, enhancing retrieval accuracy and response reliability.
- Designed and implemented a robust long-term memory framework for conversational agents utilizing (LLM-OS) capabilities. The framework focused on enabling conversational agents to store, organize, and retrieve relevant information across extended interactions, significantly improving their ability to provide context-aware and personalized responses.

Research Engineer

Dec 2015 – Dec 2016

Department of Biomedical Physics and Technology, University of Dhaka, Dhaka, Bangladesh

- Designed learning algorithms to automatically classify QRS complexes, enabling the detection of acceptable ECG traces.
- Developed PC-based telemedicine system with integrated diagnostics, enabling 500 rural families to access real-time medical services online.
- Worked on software development for automatic frequency domain analysis to identify neurological disorders from evoked EMG responses.
- Engineered software interfacing for a 12-lead computerized ECG machine on Android, facilitating near real-time data transmission to and from remote areas.

TECHNICAL SKILLS

Programming languages – Python, SQL, Matlab, Java, C, Julia.

Frameworks – PyTorch, LangChain, HuggingFace Transformers, Scikit-Learn.

ML Domains – GenAI, NLP, Time-series Forecasting, Reinforcement Learning.

Systems – Linux, OSX.