

Md Mohaiminul Islam

📍 Edmonton, AB, Canada | 🏠 Permanent Resident | 📞 +1 825-888-4849 | ✉️ mohaiminul2810@gmail.com | in linkedin.com/in/md-mohaiminul-islam

PROFILE

Senior Data Scientist with 8+ years of applied ML/NLP experience building production AI systems for regulated claims and healthcare-adjacent workflows. Recent work includes large-scale document PII detection, an employee-facing RAG application, LLM/RAG evaluation, Databricks-based release controls, model monitoring, and AI roadmap contribution. Combines hands-on ML engineering with stakeholder communication, mentorship, and hiring support.

SELECTED IMPACT

8+ yrs applied ML/NLP, research to production	12 publications, 200+ citations	Holdout human-labeled evaluation behind every model claim
---	---	---

PROFESSIONAL EXPERIENCE

Senior Data Scientist

WCB-Alberta | Edmonton, AB

Aug 2023 - Present

- Designed and operationalized end-to-end PII detection across a large historical claims archive and a continuous stream of incoming documents, combining prompt-routed LLM inference, deterministic validation, post-processing, and reviewable evidence; held quality to human-reviewed holdout labels with per-class error analysis.
- Established governed release controls for ML and LLM workloads: versioned model registry, audit logging, release gates, token/cost tracking, and developer alerts, so every promotion is reviewable and reversible.
- Built an agentic maintenance harness on hosted LLM endpoints that watches production pipelines, diagnoses failures, validates candidate fixes against sandboxed data samples, and stages deploy-ready changes for human approval; later extended it to automated code review.
- Built an employee-facing helpdesk assistant for support knowledge retrieval, pairing vector search over internal documentation with an LLM answering layer, validated-answer capture, and feedback collection; trialled it with the helpdesk team, whose answer-quality judgements drove retrieval and prompt fixes.
- Improved RAG quality with HyDE, multi-query retrieval, reranking, validated-answer capture, human-verified eval sets, LLM-as-judge scoring, and prompt/retrieval error analysis.
- Prototyped medical report-generation agents supporting extraction, validation, pre-fill recommendations, routing, human review, and audit tracking; back-tested the workflow on historical cases to size its quality and turnaround benefits before recommending adoption.
- Put a 3-class XGBoost claim-duration model into production after benchmarking NLP/Transformer alternatives on accuracy, latency, and cost, keeping the simpler model where added accuracy did not justify serving cost; shipped it with scheduled monitoring and retraining alerts.
- Led applied-AI R&D and POC initiatives for regulated claim workflows, spanning use-case scoping, hands-on prototyping, evaluation, and coordination of assigned contributors; presented recommendations to CTO and director-level stakeholders.
- Acted as the ML point person for enterprise AI adoption; contributed to AI roadmap and release-gate decisions, managed 2 direct-report interns, mentored 5 junior data scientists during project delivery, and supported 5 hiring/interview processes.

Senior Data Scientist

Quantolio | Montreal, QC

May 2023 - Jul 2023

- Prototyped a client-oriented Vision Transformer chest X-ray classifier using public medical imaging datasets; achieved ~96% accuracy and presented internally for handoff.
- Refactored a Python ML codebase and developed a Streamlit-based portfolio analytics prototype, including time-series analysis and reinforcement-learning experiments for portfolio optimization.

Lead Data Scientist, ML/NLP

Blue Guardian Canada Inc. | Toronto, ON

Feb 2023 - May 2023

- Managed 4 direct-report data science interns to deliver a fine-grained, multi-class mental-health sentiment classifier using Transformers and LLM-generated synthetic text.
- Set up an AWS Lambda-based workflow for LLM data generation, validation, and annotation that removed most of the project's manual labeling effort

CORE EXPERTISE

APPLIED AI / GENAI	ML & MODEL QUALITY	DATA / MLOPS / CLOUD	AI ENGINEERING & DELIVERY
LLMs, RAG, AI agents LlamaIndex, LangChain Databricks Apps, Vector Search HyDE, multi-query, reranking Hugging Face, vLLM	XGBoost, PyTorch, Transformers (BERT) clustering, unsupervised ML reinforcement learning LLM-as-judge, holdout evals human-verified validation error analysis, SHAP	Python, SQL, PySpark Databricks Jobs, MLflow Unity Catalog registry Docker, model monitoring retraining alerts Azure Databricks, AWS/GCP	FastAPI, model serving release gates, audit logs token/cost tracking AI roadmap planning executive technical communication mentoring, hiring support

EARLIER APPLIED ML EXPERIENCE

Applied ML Research Assistant	Sep 2015 - Dec 2023
University of Manitoba Winnipeg, MB	
- Modeled high-dimensional prediction with a privacy-preserving deep learning approach; achieved AUROC 0.94 on ~11K patient RNA-seq samples. - Created a PyTorch GNN model for protein-ligand binding scoring; achieved RMSE 1.24. - Applied XGBoost and SHAP to high-dimensional proteomics data for interpretable classification; achieved AUROC 0.83.	
Data Scientist Intern, Mitacs	Jul 2021 - Jun 2022
Servier Canada Inc. Laval, QC	
- Automated config-driven preprocessing and embedding pipelines for large SMILES molecular datasets using PySpark, Pandas, and NumPy, cutting model-preparation time across experiments. - Trained a Reformer-based generative sequence model, evaluated on chemical validity and desired-property rate.	
Data Scientist Intern	May 2019 - Sep 2019
Sightline Innovation Inc. Winnipeg, MB	
Cleaned and preprocessed graph relationship and risk-gene records in R, then implemented a maximum-flow candidate-prioritization workflow for downstream analysis.	

ADDITIONAL EXPERIENCE

Lecturer, Computer Science	May 2014 - Jul 2015
Green University of Bangladesh Dhaka, Bangladesh	
Taught 5 undergraduate computer science courses to 500+ students and coordinated 2 departmental semesters across 10+ courses and faculty schedules.	

EDUCATION

PhD Candidate, Computer Science (candidacy completed; left for industry)	2018 - 2023
University of Manitoba, Canada University of Manitoba Graduate Fellowship	
M.Sc., Computer Science	2017
University of Manitoba, Canada Manitoba Graduate Scholarship	
B.Sc., Computer Science and Engineering	2014
University of Chittagong, Bangladesh Ranked 1st in class; Prime Minister Gold Medal	

SELECTED RECOGNITION

Awards	Prime Minister Gold Medal; VADA Big Data Challenge Winner; Manitoba Graduate Scholarship; University of Manitoba Graduate Fellowship; Research Completion Award.
Publications	12 applied AI/ML publications with 200+ citations across deep learning for omics, privacy-preserving ML, and drug discovery, including a first-author paper cited 100+ times (Google Scholar, August 2026).