Ilia Alenabi

📱 +1 (778) 708-2776 | @ ialenabi@uwaterloo.ca | 🔗 linkedin.com/in/ilia-alenabi/ | 🌐 iliall.github.io

EDUCATION

University of Waterloo — Waterloo, Ontario
*Honours **Computer Science**; **Artificial Intelligence** Specialization; **Statistics** Minor* — *Sep 2022 – May 2027*

SKILLS

Languages: Java, C/C++, Python, JavaScript, TypeScript, Ruby, Golang, Rust

Frameworks/Libraries: Numpy, Pandas, PyTorch, Tensorflow, React, Django, react, Next.js

Tools/Platforms: Git, Docker, GBQ, AWS, Kubernetes, Redis, PostgreSQL, Pinecone, MongoDB

EXPERIENCE

Cerebras — Toronto, ON
Mahcine Learning Engineer - Internship — *Sep 2025 – Dec 2025*
- Integrating **foundation models** into the Cerebras **inference** system with a focus on **performance optimization**

Huawei Canada — Toronto, ON
Compiler Engineer - Internship — *Jan 2025 – Apr 2025*
- Developed an **MLIR** pass to annotate attention layers with **sharding** metadata, optimizing **tensor distribution**
- Designed an **LLVM** pass for automated software cache creation, tuning memory usage in **distributed systems**
- Analyzed **Redis**'s performance, identified hotspots, and implemented **prefetching** to reduce runtime by **20%**

Cohere — Toronto, ON
Senior Data Quality Specialist - Part-time contract — *Oct 2024 – Present*
- Evaluated AI models' problem-solving skills, solving and correcting advanced **math** and **programming** questions

Questrade Financial Group. — Remote
Software Engineer - Internship — *Sep 2024 – Dec 2024*
- Automated **cloud-based MLOps** pipeline for **fine-tuning** in **Java**, reducing redundant AI inference costs by **8%**
- Integrated **NLP** solutions—fine-tuned on **1000+** support tickets—into **DevOps**, supporting **200+** developers
- Integrated **authentication checks** into internal pipeline, auditing **SQL** code from **200+** developers for security

Silverberry Group — Vancouver, BC
Data Scientist- Internship — *Apr 2023 – Apr 2024*
- Maintained **200+** hours of audio/video in **Pinecone** vector stores utilizing **OpenAI** Whisper and **Hugging Face**
- Implemented a **U-Net** for diabetic wound segmentation and a classifier in **PyTorch**, achieving **92%** accuracy
- Developed an AI-driven **interactive agent** environment to simulate application UI behavior pre-production

RESEARCH

Vector Institute — Toronto, ON
Machine Learning Research Intern – [CompLING Lab] — *Apr 2025 – Present*
- Analyzed **vision-language models**' natural language understanding, identifying biases in **cross-modal** reasoning
- Contributed to [embedding extraction tool] for VLM models, adding support for models **Molmo** and **GLaMM**

Pingoo.AI — Remote
Researcher – [Building Trustworthy Generative AI for Diabetes Care] — *Jan 2023 – Apr 2023*
- Published in **Journal of Diabetes Science and Technology**, ranked in the **top 10%** of health sciences journals
- Reduced factual errors in AI output from **45%** to **1%** through **few-shot learning** and **prompt engineering**

PROJECTS

UW Course Planner | 🔗 | 🔗
- Developed a course schedule tool using **Django/Next.js** and integrated **OpenSSL** to secure user authentication

PQTree library | 🔗
- Optimized SageMath's [PQTree] using **Rust** with multi-threading, increasing performance by **200%** in benchmarks

UW Course Compass | 🔗
- Built a course recommendation engine in **GoLang**, improving course selection speed and efficiency for students