Aiwei (Alwyn) Yin

📱+1-647-821-2101 ✉ aiwei.yin@mail.utoronto.ca ⌨ AlwynYin 💼 aiwei-yin

EDUCATION

University of Toronto Sept. 2022 — Present
Honours Bachelor of Science Toronto, ON

- Double Major in **Computer Science** and **Statistics**; **Dean's List Scholar**. GPA: 3.67 (cumulative)
- *Selected coursework:* Operating Systems (95), Neural Networks and Deep Learning (95), Computer Graphics (95), Algorithm Design and Analysis (91), Statistical Learning (85)

EXPERIENCE

Research Assistant May 2025 — Dec. 2025
Far Data Lab, supervised by Prof. Qizhen Zhang Toronto, ON

- Conducting research on optimizing Apache Sparks task scheduler for large-scale data processing workloads.
- Performed a comprehensive literature review of task-aware scheduling in modern DBMSs, identifying key limitations in Apache Spark's default scheduling algorithms.
- Designed a query-structure-aware scheduling algorithm to improve cluster resource utilization; verified the algorithm's optimality through benchmarks and simulations.

DevOps Engineer May 2025 — Dec. 2025
Ontario Ministry of Education Toronto, ON

- Coordinated the release process for data collection systems serving educational institutions across Ontario.
- Designed, implemented, and maintained CI/CD pipelines with modern DevOps tooling, reducing manual deployment effort and improving release reliability and speed.
- Collaborated with developers and analysts to troubleshoot production incidents, streamline deployment workflows, and improve observability and monitoring practices.

PROJECTS

Self-evolving Agent System for Computational Chemistry Tasks LLM Agents

- Work conducted in the Matter Lab, supervised by Prof. Alán Aspuru-Guzik.
- Developed a self-evolving multi-agent framework for computational chemistry that automatically generates and **reuses** tools as it solves new tasks, a first-of-its-kind approach in this domain.
- Built an agent-driven tool-generation pipeline integrating scientific Python ecosystems; incorporated retrieval-augmented documentation search, code synthesis, and test-driven self-refinement to ensure high-quality tool creation.
- Preparing a manuscript based on this system, with plans to submit to *Nature Methods*.

BitNet and Quantization-Aware Training Analysis HuggingFace, PyTorch

- Investigated Microsoft's BitNet architecture to evaluate how low-precision models can scale.
- Identified critical gradient vanishing phenomena that increase training instability and cost by analyzing learned weight matrices and activation distributions.
- Proposed the need for accurate gradient estimation strategies to stabilize training for **fast-inference LLMs**.

Pintos C, x86 Assembly

Matrix Prediction for Student Performance PyTorch, NumPy

- Applied neural networks, collaborative filtering, and Item Response Theory models to predict students' performance on diagnostic questions from the *NeurIPS 2020 Education Challenge* dataset.
- Led team meetings on **strategy**, **model design**, and **progress tracking**.
- Incorporated metadata, word embeddings, and batch normalization into the model architecture, reducing false positive predictions by 9% compared to the baseline model.

SKILLS

- **Programming Languages**: Python, Java, Go, C, C++, CUDA, JavaScript, SQL, Bash
- **Data & ML Libraries**: PyTorch, vLLM, Triton, NumPy, Pandas, Matplotlib, Scikit-Learn, Apache Spark
- **Web & Backend**: React, Node.js, Express.js, MongoDB, Django
- **Tools & Platforms**: Git, Linux, CMake, AWS, GDB, LaTeX