Joseph Lee

647-262-3447 | joseph.lee1@uwaterloo.ca | www.linkedin.com/in/josephjl31 | github.com/jlee31

Education

University of Waterloo — Waterloo, ON
Bachelor of Computer Science, Co-op — *Sep 2024 – Apr 2029*
- **President's Scholarship of Distinction**

Experience

Acceleration Consortium, University of Toronto — Sep 2026 – Apr 2027
Software Developer (Full-Stack and Machine Learning) — *Toronto, ON*
- Architected a parallelized **batch simulation pipeline**, efficiently simulating hundreds of formulation predictions from the ML model, cutting compute costs by ∼**95%** and simulation runtime by **80%**
- Turned a research **Bayesian optimization** notebook into a production full-stack platform used by **global industry formulation partners**, serving ML models through a containerized **FastAPI** and **Docker** REST API
- Sped up research iteration by **50%** through the new platform, with projected savings of **$1,000+ in AWS costs** after deployment by replacing always-on notebook compute with optimized on-demand API inference
- Built an **exploratory data analysis** pipeline with published research methods (SHAP interpretability), surfacing key formulation drivers through interactive visualizations for **upcoming research publications**
- Provisioned **AWS** infrastructure as code with **Terraform** and built **GitHub Actions** CI/CD pipelines with automated validation, optimizing for **cloud portability** across AWS, **Google Cloud**, and on-premise servers

Environment and Climate Change Canada — Sep. 2025 – Dec. 2025
Software Developer in Test — *Toronto, ON*
- Engineered automation frameworks in **Python**, **Robot Framework**, **Playwright**, and **RPA Framework** across **100+ national applications**, cutting manual test cycles by **60%** and surfacing **20+ defects** pre-release
- Built **Azure SQL** validation pipelines verifying backend data, **REST APIs**, and large datasets feeding analytics and **ML workflows**, catching schema drift and data integrity issues pre-production
- Shipped **Azure DevOps** dashboards tracking defects, performance, and regression coverage, driving defect triage that lowered duplicate report rate by **40%** across the team

Environment and Climate Change Canada — May 2025 – Dec. 2025
IT Technician and Support Analyst — *Toronto, ON*
- Delivered Tier 1/2 support for **200+ end users** across **Windows**, **Linux**, **M365**, and Outlook server environments, cutting average resolution time by **50%** through reusable diagnostic playbooks
- Engineered **Python**, **PowerShell**, **Power Automate**, and **SAP GUI** scripts and automations to consolidate repetitive cross-system admin tasks, shortening task duration by **80%** across daily IT operations

Projects

Quote Finder | *Python, LangChain, FastAPI, Anthropic Claude API, React, Railway*
- Built a full-stack **RAG** application retrieving contextually relevant quotes via **vector embeddings** and **cosine similarity**, served through a **FastAPI** backend and **React** frontend deployed on **Railway**
- Developed an **agentic** workflow with **LangChain** and the **Anthropic Claude API**, where an agent interprets user queries and orchestrates retrieval tool calls to return the most relevant quotes
- Preprocessed and embedded **25,000+** quotes into a **vector store**, tuning chunking and top-k retrieval to improve result relevance and reduce query latency by **50%**

Image Processing Pipeline & SiLA 2 Connector | *PyTorch, OpenCV, SiLA 2, gRPC* — Oct 2026
SiLA Secure Lab Automation Hackathon – BEARS, Ames National Laboratory & Ursa Labs — *Awarded Distinction*
- Earned a **Distinction award** across **12 international host sites**, against PhD and Master's researchers, by building a **SiLA 2** connector for a **$40,000 Photron High-Speed Camera**
- Wrapped the camera SDK in SiLA 2 commands over **gRPC** (recording, shutter speed, memory, live view) and built an **OpenCV** pipeline turning raw frames into **ML**-ready data, with **real-time contrast enhancement**

Technical Skills

Languages: Python, TypeScript/JavaScript, C/C++, Go, Java, SQL
ML/AI: PyTorch, PyTorch Geometric, BoTorch, scikit-learn, Optuna, MLflow, LangChain, SHAP, Pandas, NumPy
Frameworks & Cloud: AWS (ECS, RDS, S3, Terraform, Cognito), Supabase, Docker, Git, FastAPI, Next.js/React