

# Shawn Gabriel Intong

✉ ishawngabriel@gmail.com    📞 +1-236-983-6065    🌐 GitHub    in Shawn Gabriel Intong

## Education

---

**University of British Columbia**

*Honours in Geophysics with Minor in Data Science*

Vancouver, BC

*Graduating May 2027*

**Coursework:** Applied ML, Numerical Approximation & Discretization, Matrix Algebra, Differential Equations II.

## Work Experience

---

**Quantitative Research Analyst**

*Legacy Bridge Capital*

February 2026 - Present

*Remote, Indonesia*

- Developing the statistical logic and variable weighting for proprietary screeners, ensuring all frameworks are grounded.
- Formulate investment-relevant hypotheses and build the statistical logic underlying the firm's scoring and ranking systems.

**Machine Learning Engineer Intern**

*Ruangguru*

June 2025 – August 2025

*Jakarta, Indonesia*

- Evaluated multiple LLMs for embedding and semantic search over a product catalogue, integrating them with Elasticsearch and comparing retrieval quality, latency, and cost trade-offs, resolving 5+ integration bugs.
- Deployed multiple ONNX-based computer vision models on GCP and evaluated inference configurations, measuring latency, accuracy, and failure cases to support scalable real-time proctoring across 10,000+ concurrent test-takers.
- Built and iterated on automated batch inference pipelines in TypeScript and Python using an internal job runner, experimenting with scheduling and batching strategies to improve throughput by 2.5× for large-volume exam workloads.
- Built an LLM evaluation and control system deployed on AWS-based infrastructure, exposing REST-style inference and evaluation endpoints via FastAPI to support scalable and repeatable model testing.

## Projects

---

**Automated Car Damage Detection with Solana Encryption** | *PyTorch, YOLO, OpenCV, Solana Devnet*

- Developed FastAPI-based vehicle damage estimator using YOLO computer vision models and Gemini AI, automating part-level damage detection and BC repair cost calculation with pre/post-tax totals in CAD.
- Implemented two-stage ML inference pipeline with severity classification algorithms and integrated Solana blockchain attestation for tamper-evident damage report verification and audit trails.
- Built production-ready REST API with 8+ endpoints, dual UI/JSON interfaces, pytest test coverage, and flexible deployment modes supporting mock, real-time, and auto-detection inference strategies.

**LLM Guardrail Inference Pipeline** | *PyTorch, FastAPI, SQLAlchemy, MCP Agents*

- Built and iterated on a PyTorch-based LLM guardrail system to evaluate model behavior, including toxicity detection, hallucination similarity checks, exposing simple inference and evaluation endpoints through FastAPI.
- Experimented with agent MCP-style tool-calling to safely connect LLMs to external tools such as a UTC time function, debugging unsafe behaviors and comparing outcomes across different prompt and tool configurations.
- Designed a lightweight telemetry and evaluation pipeline using SQLAlchemy to log prompts, model outputs, guardrail triggers, similarity scores, and latency, enabling experiment tracking, debugging, and future fine-tuning.

**3D-Consistent Diffusion-Based Image Generation** | *PyTorch, Diffusers, ControlNet, LoRA, Open3D*

- Designed and implemented an end-to-end geometry-aware diffusion pipeline, converting 3D assets into multi-view depth maps and conditioning Stable Diffusion, to better understand model behavior under geometric constraints.
- Spearheaded a custom data generation stack using Open3D to render point clouds into aligned depth views of 5 objects.
- Ran controlled experiments comparing baseline and LoRA-fine-tuned models, tracking depth error metrics and diagnosing failure cases, obtaining a mean depth L1 error of 0.5424, decreasing error by 9.34% relative to the baseline.

## Awards & Achievements

---

**1st Place of UBC Data Science Club Datathon 2025**

- Used N-BEATS & PatchTST time series forecasting models to forecast emerging and declining industries in Vancouver, achieving over a 70% reduction in MSE compared to using traditional ML models such as XGBoost & LightGBM.

**2nd Place of UBC Data Science Club x DarkVision Computer Vision Hackathon 2024**

- Reimplemented the 2023 paper on point cloud extraction to ensure increased accuracy in measurements crucial for NDT.

## Summary of Qualifications

---

**Languages:** Python, TypeScript, R.

**Frameworks & Tools:** PyTorch, TensorFlow, FastAPI, HuggingFace Transformers, Git, Bash, Docker, OpenCV, Gradio, Google Cloud Platform, Google Colab, TorchProfiler.

**Libraries:** NumPy, Pandas, SciPy, sentence-transformers, FAISS, spaCy, scikit-learn, NLTK, HuggingFace Datasets.