RHYS DAVID MIRANDA

Hong Kong • rhysdmiranda@gmail.com • rhysmiranda.github.io
github.com/RhysMiranda • linkedin.com/in/rhys-miranda-760a71279

EDUCATION

City University of Hong Kong — Hong Kong
BSc (Hons) Data and Systems Engineering, Advanced Standing I — Sep 2023 – Feb 2027 (expected)
Relevant coursework: Fundamentals of Machine Learning, Statistical Methods and Data Analysis, System Modelling and Simulation, Data Intensive Computing, Financial Engineering and Analytics, Decision Analytics and Risk Management.

University of Auckland — Auckland, New Zealand
Exchange programme in cybersecurity and artificial intelligence — Jul – Nov 2025

Hong Kong Metropolitan University — Hong Kong
BSc (Hons) Computer Science. cGPA 3.88/4.00, Dean's List — Sep 2022 – Jun 2023
Transferred to City University of Hong Kong with Advanced Standing.

EXPERIENCE

NVIDIA — Hong Kong
Student Ambassador; co-researcher with the NVIDIA AI Technology Centre (NVAITC) — Feb 2026 – Present

- Built a 3-agent LLM service (retrieval, fraud detection, compliance) over a real-time streaming feed for SME lending, validating every agent output against a schema so a malformed or unsupported judgment is rejected before anything acts on it.
- Built the team's risk data system from scratch, resolving 16 source systems in 13 formats into one record per entity with 30+ derived risk indicators on 1M+ rows, and moved it to GPU (RAPIDS cuDF, cuML) for roughly an order-of-magnitude speed-up.
- Built the evaluation framework the team measures model changes against: one fixed benchmark, three replicated runs, and results cached on a content hash so any earlier run reproduces exactly.
- Attached SHAP attributions to every scored record so each decision can be explained to the human reviewer who acts on it, and wrote the documentation the system is operated from.

Guns and Rain — South Africa (remote)
Data and AI Specialist (Internship) — Jun – Aug 2025

- Built and deployed an LLM information-extraction tool that turns unstructured newsletter and social media text into structured records against a taxonomy I designed, replacing a manual weekly review.
- Migrated and standardised 2 relational database systems with validation rules and data dictionaries, cutting data retrieval time by 40% for 15+ stakeholders.
- Presented findings to non-technical clients every week, turning model output into recommendations they could act on.

Just Grace NPC — South Africa
Data Scientist (Internship) — Jun – Aug 2024

- Integrated 4 cloud data sources into a Power BI reporting layer, with dashboards tracking Key Risk Indicators for senior management.
- Designed and delivered data management and Power BI training for 20+ staff and managers.

Suca Foods — Buon Ma Thuot, Vietnam
Website Developer (AIESEC Global Volunteer) — Jun – Aug 2024

- Led development of the company's websites, improving how its products and services are presented to customers.

SELECTED PROJECTS

silver-signal: News, Overnight Moves and Silver ETF Pricing — 2026
PySpark, FinBERT, BigQuery, Airflow, MongoDB, FastAPI, React, Docker, R · github.com/RhysMiranda/silver-signal

- Scored 233,737 GDELT headlines with FinBERT in PySpark and showed under walk-forward evaluation that news adds nothing beyond the overnight silver and rupee move; found instead that the ETF opened rich and gave back 17.9 bps a day in its first hour ($t = -4.0$; Newey-West and block bootstrap in R), an edge that faded to zero in 2026.
- Built a daily Airflow pipeline in Docker that paper-trades three strategies into MongoDB Atlas, served by FastAPI on Vercel and a React dashboard on GitHub Pages.

Synthetic Graph Generator for SME Loan Fraud — 2026
Python, PyTorch Geometric, Bank Account Fraud dataset (NeurIPS 2022)

- Built a generator that embeds organised fraud rings as network structure rather than feature values, because labelled fraud data is private, so a graph model can be tested against a non-graph model on identical information.

- Reported every result as lift over a gradient-boosted baseline under leakage-free grouped splits, averaged over 3 seeds. At a realistic 1% fraud rate the graph model reached 0.528 PR-AUC against 0.193, recovering 2.7 times more fraud and catching two fraud categories the baseline missed entirely.

Neural and LLM-Guided Decoders for Quantum Error Correction 2025

City University of Hong Kong. PyTorch (CUDA), Stim, PyMatching, Google quantum hardware data

- Trained autoregressive neural decoders on GPU, a masked autoencoder (MADE) and a transformer density estimator (TraDE), conditioned on the error syndrome, and benchmarked them against BP-OSD.
- Built an optimiser in which an LLM proposes edge-weight changes to a minimum-weight perfect matching decoder, accepted only when held-out error improves; compared 7 strategies against dual annealing and TuRBO on 45,000 held-out shots.
- Reduced logical error rate by 2.8% on simulated surface-code data and by 0.74% on real Google quantum hardware data, reported separately.

Limit Order Book Matching Engine and Transaction Cost Analysis

C++17, CMake, no third-party dependencies

- Built a price-time priority matching engine that replays 1,000,000 messages into 612,549 trades at 148.9 ns per message (6.71M messages per second), passing 54 assertions with 0 failures.
- Chose each data structure against its access pattern (ordered maps for price levels, intrusive lists for safe mid-queue cancels, integer ticks instead of floating-point keys) and added a TCA layer measuring orders against arrival price, VWAP, TWAP and close.

LLM Serving Stack with Schema-Enforced Output

Python, Hugging Face Transformers, gRPC, FastAPI, Pydantic

- Built an internal gRPC service behind a public FastAPI gateway with server-sent-event streaming, serving a local Qwen3-0.6B model fully offline, and enforced a JSON schema on every response so malformed output becomes a typed defect record instead of an answer.

S&P 500 Forecasting and Portfolio Optimisation

TensorFlow/Keras, scikit-learn, cvxopt

- Forecast the index from 2000 to 2023 with an LSTM against support vector regression and random forest baselines, then fed the forecasts into a convex mean-variance optimisation to produce portfolio weights.

LEADERSHIP AND AWARDS

HKSTP One Person Company (OPC) Hackathon Hong Kong

Winner Aug 2026

- Built a full-stack AI debate and speech coach that pairs browser-based voice and eye tracking with an AI scoring argument quality across four debate criteria.

HKSTP Talent Foundry Hong Kong

Participant Oct 2025 – Present

- 100+ hour applied programme with NVIDIA, AWS and Deloitte; nominated as a student speaker at the HKSTP career fair and the OPC Hackathon.

AIESEC in CityU Hong Kong

Local Committee Vice President, Outgoing Global Volunteer Sep 2023 – Jun 2025

- Led a 7-person team whose campus campaign sent 14 students on volunteer exchanges abroad.

HK Techathon+ 2024 Hong Kong

Merit Award Nov 2023 – Jan 2024

- Designed an add-on device giving wheelchairs autonomous stair-climbing, with energy recovery as a backup power source.

SKILLS AND INTERESTS

Programming: Python (daily), C++17, R, Scala, SQL, JavaScript/TypeScript, Node.js, React, Bash

Machine learning and LLMs: PyTorch, PyTorch Geometric, TensorFlow/Keras, scikit-learn, XGBoost, SHAP; graph neural networks, density estimation, reinforcement learning; multi-agent systems, RAG, LangChain, Hugging Face Transformers, NVIDIA NIM

Big data and databases: Hadoop, Apache Spark, Apache Flink, Apache Kafka, Apache Airflow, Pig Latin, HiveQL; RAPIDS cuDF, cuML, cuGraph; MongoDB and other NoSQL stores, relational databases

Engineering and tools: Git, Docker, FastAPI, gRPC, Flask, AWS, Triton Inference Server, Tableau, Power BI

Statistics and quant: hypothesis testing, bootstrap inference, Monte Carlo simulation, convex optimisation, backtesting, transaction cost analysis

Certifications: AWS Certified AI Practitioner; AWS Certified Cloud Practitioner; NVIDIA DLI: Building Transformer-Based NLP Applications, Building LLM Applications with Prompt Engineering

Languages: English (native; IELTS Academic 8.0), Mandarin (studying toward HSK 3)

Interests: basketball, table tennis, reading, travel