Martin Ho

MSc Applied Data Science

@ martin.ho.zy97@gmail.com | [redacted]

linkedin.com/in/martin-ho-zy | github.com/minimartzz

SUMMARY

Results-driven Data Scientist with 3 years of experience architecting scalable data solutions across the full ML lifecycle. Proven track record in streamlining data ingestion, designing predictive models, and deploying production-grade ML solutions. Expert at translating complex datasets into actionable business intelligence to optimise operational efficiency and ROI. Specialised in Python, SQL, and others in the data science toolkit to build innovative, user-centric data products.

EXPERIENCE

Nxera — Oct 2023 - Nov 2024

Data Engineer

- Engineered and scaled a robust data lake and ETL pipeline architecture using Kafka, Python and MinIO automating the ingestion of 300k+ daily data points
- Developed a real-time monitoring suite using in-house tools to track machine health and reduce downtime
- Led the integration of AI-driven optimisation algorithms within data center operations achieving S$300k in annual cost savings

Singtel — Oct 2022 - Sep 2023

Machine Learning Engineer (Management Associate Rotation)

- Architected a Time-Series Forecasting model (ARIMA) to predict spare fiber utilisation, identifying high-potential sales territories and increasing pipeline conversion rates by 23%
- Optimised the enterprise quotation engine by implementing a rule-based engine slashing lead times by 20%
- Standardized the ML lifecycle by building and internal MLOps platform using Mlflow, Ansible, Bamboo, streamlining model versioning and deployment to boost team productivity by 10%

NCS — Oct 2021 - Sep 2022

5G Consultant (Management Associate Rotation)

- Architected cloud-native 5G proof-of-concepts on Azure, delivering technical demonstrations to client stakeholders that secured multiple new enterprise partnerships
- Consulted on digital transformation strategies for cross-industry clients, designing 5G-enabled products that expanded the firm's addressable market in Retail and Manufacturing
- Cultivated a strategic ecosystem of 50+ technology partners, facilitating collaborative R&D efforts to create new offerings for emerging 5G solutions

EDUCATION

Masters of Science (MSc) in Applied Data Science: 4.0 GPA (projected), *University of Michigan* — Apr 2026

Bachelors of Engineering (BEng) Business Analytics, AI Minor: Honors w/ Distinction, *SUTD* — Jun 2021

SKILLS

AI & Machine Learning	Supervised/ Unsupervised Learning, LLMs, Agents
Data Engineering	ETL/ELT Pipelines, Data Lakehouse Architecture, Schema Design, Real-time Data Ingestion, Deployment and Management
Language & Frameworks	Python (Pandas, Numpy, Sklearn, PyTorch, TF), SQL (PostgreSQL, MySQL), FastAPI, HuggingFace, ChromaDB, Langchain
Web Development	NEXT.js, TailwindCSS, Supabase
Cloud & MLOps	AWS, Mlflow, Docker, Ubuntu, Git
Visualisation & BI	Grafana, Tableau, Matplotlib/ Seaborn, Adobe Illustrator

CERTIFICATIONS

Microsoft Azure Developer Associate *(Microsoft)*

HuggingFace LLM Certificate *(HuggingFace)*

Web Development Bootcamp *(Udemy)*

Postgraduate Diploma in Business Management *(Emeritus)*

AWS Solutions Architect Associate *(AWS) in progress*

NOTABLE PROJECTS

Trakka

Solo-developed an end-to-end web application for board game analytics. Enabling users to track games, and compare performances providing insight into their strategies

Features

- Sessions tracking and performance metrics
- Group and role-based management
- Real-time notification system

Tech Stack: NEXT, Supabase, Tailwind

pAInt-by-the-numbers

Developed an AI pipeline that integrates Diffusion models with Computer Vision algorithms to transform standard imagery into "paint-by-the-numbers" templates

Features

- Number of colour selection
- Auto-colour number assignment to grid
- OpenAI integration

Tech Stack: Python, OpenAI, sklearn, opencv, Docker

Kaggle Competitions

Participated in a variety of Kaggle competitions to learn and practice data science skills

Competitions

- Charting Student Math Misunderstandings
- Crypto Market Prediction
- Predicting Podcast Listening Times
- Mechanism of Action Prediction

Tech Stack: Python

FHIR Medical Cost Prediction

Developed a prediction model that estimates the cost of a patients US hospital journey based on grouped diagnositics

Features

- Unsupervised learning for bundled diagnoses
- Supervised learning for cost prediction
- Breakdown of cost contribution per diagnoses

Tech Stack: Python, sklearn, XGBoost