

Results-driven Software Engineer and Computer Science undergraduate with a data engineering focus.

Experience

• Inland Revenue Authority of Singapore (IRAS)

Data Engineer Intern

Jun 2025 - Jan 2026

- Modernized refund risk data workflows by converting legacy SAS processing logic to SQL, streamlining the organization's data stack and improving maintainability.
- Designed and implemented a Jinja-based transpiler enabling shared data processing logic to run seamlessly on both Spark and SQLite, supporting tailored training and inference needs.
- Architected a multi-stage data pipeline with Kedro, Azure ML, and Azure Synapse to build a LanceDB vector database for an internal Retrieval-Augmented Generation (RAG) application, enhancing data accessibility and performance.
- Automated vector database refresh processes, eliminating manual operations and saving engineering hours, enabling faster and more reliable data updates.

• Gojek

Machine Learning Engineer Intern

Mar 2020 - Oct 2020

- Engineered an open-source Machine Learning (ML) Feature Store for Gojek's driver-customer matchmaking model, enhancing data processing efficiency using Python, Java, Golang, Kafka, and Redis.
- Resolved critical production issue by correcting Redis storage schema mismatch, ensuring uninterrupted matchmaking model performance.
- Designed and deployed a comprehensive analytics dashboard using Grafana and GCP BigQuery, empowering management with real-time insights for data-driven decision-making.

Projects

• Flowmotion

<https://mrzzy.co/work/flowmotion-real-time-routing-for-sg-drivers>

Sep 2024 - Nov 2024

<https://github.com/ZhiXin18/flowmotion>

- Engineered a real-time traffic congestion routing app that leverages live camera feeds and machine learning to deliver optimized routes for Singaporean drivers, improving commute efficiency.
- Integrated a custom YOLOv8-based machine learning model to analyze traffic congestion, delivering real-time insights for route optimization using the OSRM API.
- Developed a cross-platform mobile application with Flutter, ensuring seamless functionality on both iOS and Android, while managing scalable cloud infrastructure using Firebase and Google Cloud.

• Providence

<https://mrzzy.co/work/providence-data-engineering-personal-finance>

Nov 2023 - Nov 2024

<https://github.com/mrzzy/providence>

- Designed and automated a robust ELT pipeline that integrates personal finance data from multiple sources, enabling data-driven insights and informed budgeting decisions.
- Optimized cloud architecture by migrating to DuckDB, reducing data storage and query costs while enhancing query performance, ensuring scalability with a low operational overhead.
- Built a dynamic financial dashboard using Apache Superset, providing real-time insights into spending trends and financial health with a user-friendly interface for personal finance management.

Education

• Nanyang Technological University

Bachelor of Computing in Computer Science

Aug 2023 - Dec 2027

• Ngee Ann Polytechnic

Diploma in Information Technology

Apr 2018 - Apr 2021

Skills

- Backend: Python, SQL, Java, JavaScript, TypeScript, Node.js, Golang, Django, Jinja
- Cloud: GCP, Firebase, AWS, Azure, Terraform, Docker, Ansible, Packer, Kubernetes, CI/CD, Github Actions
- Data: PostgreSQL, Redis, BigQuery, Grafana, DBT, Prefect, Spark, DuckDB, Firestore, Data Modeling, Data Warehousing
- Other: Linux, Cloud Infrastructure, Data Analytics