Michael Hoon Yong Hau

+65 97769494 | michael_hoon@mymail.sutd.edu.sg | linkedin.com/in/michaelhoon | github.com/michael-hoon

Education

Singapore University of Technology and Design — May 2026 – Dec. 2026
Master of Engineering (Research), Artificial Intelligence — *(Expected)*

Singapore University of Technology and Design — Sep. 2022 – May 2026
Bachelor of Engineering, Engineering Systems and Design, Business Analytics Specialisation — *GPA: 4.59/5.0*

University of California, Berkeley — Jun. 2023 – Aug. 2023
Summer Exchange, Econometrics — *Berkeley, CA*

Experience

AI Engineer Intern — Apr. 2024 – Dec. 2024
DSO National Laboratories — *Singapore*

- Deployed Retrieval Augmented Generation (RAG) pipeline for internal research assistant LLM application with open-sourced tools, significantly reducing man-hours required for staff to conduct research & policy analysis.
- Implemented Advanced Agentic RAG techniques with Neo4j Knowledge Graph retrieval and Plan-and-Solve prompt engineering, significantly improving reasoning & generation capabilities compared to vanilla RAG pipelines.
- Optimised data pre-processing, generation, and retrieval pipelines through a comprehensive literature review of state of the art RAG techniques, to improve retrieval relevancy and ensure quality generated answers.
- Participated in deploying the model within the company's local IT platforms, ensuring scalability and performance.

Data Analyst Intern — Jan. 2024 – Apr. 2024
Poh Tiong Choon Logistics Ltd. — *Singapore*

- Worked on Supply Chain Analytics, analysing cross-department delivery driver dataset to derive insights and optimize business strategies. Identified key areas for improvement in scheduling and boosted driver utilisation rates by 20%.
- Developed end-to-end automated data pipeline solution for cleaning and analysis with bash scripts, integrating existing database system with Python, significantly reducing working man-hours.
- Created PowerBI Dashboard for key metrics and patterns on route analysis. Wrote internal report discussing findings

Data Analyst Intern — Sep. 2022 – Jan. 2023
Aviation Studies Institute, SUTD — *Singapore*

- Extracted key business insights for industry partner, International Air Transport Association (IATA), via Exploratory Data Analysis in R, SQL, and Python. Conducted economic analysis of air cargo demand in Southeast Asia.
- Developed correlation models for predictive analysis of internal air cargo demand data with open-sourced economic data. Evaluated hypothesis on key drivers of regional air cargo demand.
- Wrote an internal paper discussing findings, presented to IATA and the Civil Aviation Authority of Singapore.

Projects

LepakLah! | *Python, FastAPI, MySQL, Docker, Kubernetes, Flutter* — Aug. 2024 – Sep. 2024

- Developed a Flutter-based mobile application as part of Dell InnovateFest, a Public-Good hackathon to combat social isolation among the Elderly. Represented University and **won 3rd place award.**
- Integrated a generative AI model (LLM) to produce personalized workshop ideas based on existing data, and an image-generation model to create visual samples of activities. Models deployed on NVIDIA NIM platform.
- Developed System Architecture, containerised application backend with Docker and managed with Kubernetes, ensuring scalability and efficiency on a cloud-native environment. Deployed on Red Hat OpenShift platform.

CareerQuest | *React.js, Vite.js, Tailwind CSS, Firebase* — Aug. 2024 – Sep. 2024

- Built a full stack responsive web application as part of the Open Goverment Products Build for Good 2024 hackathon, a platform that supports Education and Career Guidance for Junior College students.
- Used React.js + Vite.js and Tailwind CSS for the frontend UI, with Firebase as the backend server, deployed application on Firebase Cloud. Ensured proper CI/CD practices using GitHub workflows.
- Conducted proper user research, user testing, and stakeholder analysis with relevant government statutory boards, and incorporated the entire product design workflow in designing the solution.

Transportation Analytics | *R, Discrete Choice Modelling, Econometrics* Aug. 2024 – Aug. 2024

- Internal Kaggle Competition on transportation analytics with General Motors dataset, to predict consumer choices among bundles of safety features in cars using Discrete Choice Models.
- Developed and tested multiple models in *R* such as Multinomial Logit, Mixed Logit, Random Forest, and XGBoost, validating model results on a private test set using a cross-entropy loss metric.
- Fine-tuned model with hyperparameter grid search and introduced stacking techniques such as Meta-Learning to implement a soft-voting classifier, obtaining **2nd place overall**. Discussed model interpretability with SHAP values.

MRP Backend System | *Python, SQL, Docker, Microsoft Azure, Airflow* Jun. 2024 – Aug. 2024

- Designed and implemented a backend system for Materials Requirement Planning (MRP) for a supply chain project, optimizing inventory management by forecasting material requirements and automating workflows.
- Developed an end-to-end ETL pipeline with Python, SQL, Azure Cloud Platform, and orchestrated automated DAG workflows with Apache Airflow. Back-end microservices containerised with Docker, validated data with Pydantic.
- Integrated system with SQL database to store and retrieve data on Microsoft Azure Cloud Platform, enabling a scalable, secure, and reliable cloud-based data storage.

Relevant Coursework

- Applied Machine Learning
- Optimisation
- Natural Language Processing
- Statistics & Econometrics
- Data Science
- Data Engineering

Skills

Languages: Python, R, SQL, Julia, LaTeX
Tools: Git, Docker, Kubernetes, Linux/Bash, Ollama, Neo4j, Microsoft Azure, Firebase, Apache Airflow
Libraries/Frameworks: LangChain, HuggingFace, PyTorch, Pandas, NumPy, scikit-learn, FastAPI, tidyverse (R)