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　　　　　　　　　　　　　Sep. 2022 – May 2026
Bachelor of Engineering, Engineering Systems and Design, Minor in Computer Science　　　　　*Singapore*

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

EXPERIENCE

AI Software Engineer Intern　　　　　　　　　　　　　　　　　　　　　Mar. 2024 – Dec. 2024
DSO National Laboratories　　　　　　　　　　　　　　　　　　　　　　　　*Singapore*

- Deployed Retrieval Augmented Generation (RAG) pipeline for internal research assistant chatbot application with open-sourced tools, to provide intelligent responses based on large-scale unstructured data sources.
- Implemented Advanced Agentic RAG techniques incorporating DAG workflows with intelligent Plan-and-Solve prompt engineering, significantly improving reasoning capabilities compared to vanilla RAG pipelines.
- Optimised Data Pre-processing, Generation, and Retrieval pipelines through a comprehensive literature review of state of the art NLP techniques, to improve retrieval relevancy and ensure quality generated answers.
- Participated in deploying the model within the company's 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 and Power BI dashboard, significantly reducing working man-hours.

Data Analyst Intern　　　　　　　　　　　　　　　　　　　　　　　　Sep. 2022 – Jan. 2023
Aviation Studies Institute　　　　　　　　　　　　　　　　　　　　　　　　*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, SQL, Docker, Kubernetes, Flutter*　　　　　　　　　　　Aug. 2024 – Sep. 2024

- Developed a Flutter based mobile application as part of the Dell InnovateFest finalist, a Public-Good hackathon to combat social isolation among the Elderly.
- Implemented a recommender system to match users with activity buddies, leveraging data from user profiles and preferences to generate optimal matches.
- 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.
- Containerised application backend with Docker and managed with Kubernetes, ensuring scalability and efficiency on a cloud-native environment. Deployed on Red Hat OpenShift platform.

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

- Internal Kaggle Competition on developing Discrete Choice Models for transportation analytics with General Motors dataset, to predict consumer choices among bundles of safety features in cars.
- Developed 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.

SKILLS

Languages: Python, R, SQL, Julia, LaTeX
Tools: Git, Docker, Kubernetes, Ollama, Linux/Bash, Microsoft Azure Cloud Platform
Libraries/Frameworks: LangChain, LangGraph, HuggingFace, Pandas, NumPy, scikit-learn, FastAPI, tidyverse (R)