





# Hew Li Yang

Data Scientist | Machine Learning Engineer

 hewliyang  hewliyang

 Singapore  +65 98122803  hewliyang@u.nus.edu  <https://hewliyang.com>

## Technical Skills

### Language

Python, Typescript/Javascript, Java, R, SQL

### Platform

Kubernetes/Openshift, Docker, cron, Github Actions, Vercel, Azure OpenAI


### Library

Transformers, Torch, FastAPI, Svelte, Streamlit, Helm

## Selected Projects

### Autograd UI


Interactive tracing of computation graphs a la Torch autograd

 <https://autograd.hewliyang.com>

Classroom tool for Introduction to Machine Learning courses

### NUS Sentiment


Sentiment analysis & semantic search over the NUS subreddit

 <https://devpost.com/software/nus-sentiment>

Built in 24 hours. Winning project at NUS largest annual hackathon.

### Traffic Detection on Singaporean Highways

Leveraging computer vision and the Land Transport Authority's traffic camera feed to automate traffic detection, prediction & analytics.

 <https://hewliyang.com/blog/traffic-detection>

## Professional Experiences

### GIC

Machine Learning Engineer Intern

June 2023 - Dec 2023

Singapore

- Key contributor to an internal Python package for LLM orchestration with high utilization among developers and quants, focusing on zero-shot classification & structured data extraction
- Evangelized Elasticsearch as a hybrid full-text and vector search engine, improving performance of Retrieval Augmented Generation use cases across the board
- Led, developed and collaborated with multiple engineering functions on a high impact POC translating natural language into search queries, reducing time taken for traders to fetch large time series datasets
- Adapted software engineering best practices such as linting, type-checking, containerization with Docker and deploying Kubernetes resources with Helm Charts

### Infineon

Data Science Intern

May 2023 - June 2023

Singapore

- Streamlined data cleaning process by converting notebooks to Python modules for improved differentiability between revisions & enhancing reusability
- Fine-tuned a BERT model for text classification on internal documents with Transformers, reducing human hours required to sort them into the correct categories

### National University of Singapore

Teaching Assistant

Aug 2022 - May 2023

Singapore

- Typeset model answers for ST1131 (Introduction to Statistical Computing) exercises with R Markdown and LaTeX
- Answer student queries and facilitate discussion on probability & statistical concepts during tutorial sessions

## Education

### National University of Singapore

Data Science & Analytics, Minor in Computer Science

4.3/5

Aug 2020 - June 2024

Bachelor

**Selected coursework:** Linear Algebra, Convex Optimization, Object Oriented Programming, Big Data Systems

**Awards:** Coreteam's Best Roll at NUS Hack&Roll 2023, beating over 100 teams and 400 participants