

# Timothy Lee

timothylyhy@hotmail.com | +65 98572155

<https://timooothy.dev> | <https://linkedin.com/in/timooothy> | <https://github.com/timooo-thy>

## EDUCATION

---

<b>Bachelor of Science (B.S.) - Data Science and Artificial Intelligence</b> <i>Nanyang Technological University, Singapore</i>	Aug 2022 – Jun 2026 (Expected)
<ul style="list-style-type: none"><li>– <b>CGPA:</b> 4.75/5.00</li><li>– <b>Scholarships:</b> Nanyang Technological College Scholarship, Bigo Excellence Scholarship</li><li>– <b>Awards:</b> Dean's List AY 23/24</li><li>– <b>Hackathons:</b> SimplifyNext RPA x Gen AI 2024 (1st Place), DSTA BrainHack 2024 (1st Runner Up), GIC Code To Impact 2024 (2nd Runner Up), HTX HacX 2024 (3rd Runner up)</li><li>– <b>Exchange:</b> Singapore Management University, University of Waterloo</li><li>– <b>Certifications:</b> Google Analytics Specialisation, Google UX Specialisation</li></ul>	

<b>Diploma with Merit - Engineering Science</b> <i>Ngee Ann Polytechnic, Singapore</i>	Apr 2017 – Apr 2020
<ul style="list-style-type: none"><li>– <b>CGPA:</b> 3.99/4.00</li><li>– <b>Awards:</b> Terasaki Electric Medal, Ngee Ann Scholarship</li><li>– <b>Certifications:</b> Diploma Plus in Advanced Engineering Mathematics</li></ul>	

## EXPERIENCE

---

<b>Machine Learning Engineer Intern</b> <i>TikTok, Singapore</i>	Aug 2025 – Dec 2025
<ul style="list-style-type: none"><li>– Engineered high-precision Small Language Models (SLMs) via Knowledge Distillation to capture user intent; applied a novel “Random Retrieval Augmented Generation (RAG)” fine-tuning strategy for GeoIntent (0.90 P/R) and semantic query analysis for Timeliness (0.96 P/R).</li><li>– Designed and deployed a near-line cache architecture across 16 countries (US, EU, ROW) to serve these query intent signals in real-time.</li><li>– Integrated GeoIntent and Timeliness query intent signals into Ranking and Retrieval pipelines across 16 countries (US, EU, ROW), measurably improving local discovery relevance for location-based and time-sensitive search results.</li></ul>	

<b>Software Engineer Intern</b> <i>Goldman Sachs, Singapore</i>	May 2025 – Aug 2025
<ul style="list-style-type: none"><li>– Engineered a Policy Builder Playground for authorisation policies, featuring custom syntax highlighting and an integrated chatbot delivering 95% compilable code for over 100 internal developers.</li><li>– Developed a graph-based Self Healing Code Agent leveraging Retrieval Augmented Generation (RAG), enabling automated code correction, saving an estimated 10 hours per developer weekly.</li><li>– Architected and deployed the Self Healing Code Agent backend on Kubernetes, supporting horizontal scaling and high availability for internal developer tooling.</li></ul>	

<b>Machine Learning Researcher Intern</b> <i>Digital Trust Centre, Nanyang Technological University, Singapore</i>	Jul 2024 – Dec 2024
<ul style="list-style-type: none"><li>– Led development of secure AI systems researching and implementing ML defences against adversarial attacks via Machine Unlearning.</li><li>– Research areas include Machine Unlearning, Backdoor Detection in AI.</li></ul>	

<b>Software Engineer Intern</b> <i>ShopBack, Singapore</i>	Jul 2024 – Dec 2024
<ul style="list-style-type: none"><li>– Spearheaded implementation of Generative AI solutions to optimise internal operations, reducing average handling time (AHT) from 365 seconds to 200 seconds (45% improvement).</li><li>– Developed a prompt management version control platform integrating Salesforce and Zendesk, resulting in 80% improved consistency and 60% faster workflows.</li></ul>	

## Machine Learning Engineer Intern

Apr 2024 – Jul 2024

*Home Team Science & Technology Agency, Singapore*

- Built a multimodal Large Language Model (LLM) chatbot integrating LLaVA, Whisper, and LLaMA3, reducing case processing time by over 40% for officers.
- Implemented memory capabilities and optimised query retrieval using LangChain and Chroma via Retrieval Augmented Generation (RAG) and advanced metadata filtering.
- Finetuned a LLaMA3 large language model using QLoRA with domain knowledge, improving accuracy by 30% after evaluation.

## Full Stack Developer Intern

Dec 2023 – Mar 2024

*Siftee, Singapore*

- Engineered SEO-optimised web applications using Nuxt.js, mastering state management and dynamic data rendering to enhance user experience and interface responsiveness.
- Implemented secure cloud infrastructure using Google Cloud Platform, integrating Stripe for transactions. Optimised data operations and storage, boosting retrieval speed by 40% and transaction success rates by 25%.
- Leveraged OpenAI with Python to automate ETL tasks, boosting backend efficiency by 50% and ensuring scalable data transformation and integration workflows.

## Cloud Engineer Intern

May 2023 – Aug 2023

*Singapore Land Authority, Singapore*

- Spearheaded the Proof of Concept (POC) on Microsoft Azure and led data virtualisation initiatives utilising geospatial insights via Denodo, elevating data accessibility and fostering enriched insights.
- Designed user-centric dashboards to transform intricate data into actionable insights tailored to users' needs using MicroStrategy.

## SKILLS

---

**Languages:** Python, TypeScript, Golang, SQL

**Machine Learning & AI:** PyTorch, LangChain, LangGraph, HuggingFace Transformers

**Databases:** PostgreSQL, MySQL, MongoDB, Chroma, Redis

**Backend & Web Frameworks:** FastAPI, Next.js, Express.js, Gin

**Cloud & DevOps:** AWS, Google Cloud Platform, Docker, Kubernetes

## ACADEMIC PROJECTS

---

### Recommendation System Anomaly Detection

2024

- Competed in a machine learning competition, achieving a 94% anomaly detection AUC-ROC score by employing deep neural networks, logistic regression, and one-class SVM.
- Enhanced model performance through extensive feature engineering, utilising techniques such as SMOTE and ELU, successfully identifying and mitigating anomalies in a recommendation system.

### AI Product Chatbot

2023

- Designed a conversational AI Streamlit Python app using Langchain and Hugging Face models.
- Enhanced user experience by incorporating features like Conversational Memory, dynamic interaction, and session management, improving context awareness and user experience.
- Utilised Retrieval Augmented Generation for precise and relevant real-time FAQ responses.

### Singapore University Guide

2023

- Designed, developed, and deployed a comprehensive university guide using the MERN stack.
- Integrated modern UX techniques, responsive design, ensuring an engaging user experience.
- Seamlessly incorporated OAuth, email notifications, and OpenAI's GPT API.