

Timothy Lee

Singapore | timothylhy@hotmail.com | +65 9857 2155
timoothy.dev | linkedin.com/in/timoothy | github.com/timooo-thy

EDUCATION

Bachelor of Science (B.S.) - Data Science and Artificial Intelligence <i>Nanyang Technological University, Singapore</i>	Aug 2022 - Jun 2026 (Expected)
<ul style="list-style-type: none">CGPA: 4.75/5.00Scholarships: Nanyang Technological College Scholarship, Bigo Excellence ScholarshipAwards: Dean's List AY 23/24Hackathons: 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)Certifications: Google Analytics Specialisation, Google UX Specialisation	

EXPERIENCE

Machine Learning Engineer Intern <i>TikTok, Singapore</i>	Aug 2025 – Dec 2025
<ul style="list-style-type: none">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).Designed and deployed a near-line cache architecture across 16 countries (US, EU, ROW) to serve these query intent signals in real-time.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.	
Software Engineer Intern <i>Goldman Sachs, Singapore</i>	May 2025 – Aug 2025
<ul style="list-style-type: none">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.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.Architected and deployed the Self Healing Code Agent backend on Kubernetes, supporting horizontal scaling and high availability for internal developer tooling.	
Software Engineer Intern <i>ShopBack, Singapore</i>	Jul 2024 – Dec 2024
<ul style="list-style-type: none">Spearheaded implementation of Generative AI solutions to optimise internal operations, reducing average handling time (AHT) from 365 seconds to 200 seconds (45% improvement).Developed a prompt management version control platform integrating Salesforce and Zendesk, resulting in 80% improved consistency and 60% faster workflows.	
Machine Learning Engineer Intern <i>Home Team Science & Technology Agency, Singapore</i>	Apr 2024 – Jul 2024
<ul style="list-style-type: none">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.	

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