

Dang Khoa Le

103844421@student.swin.edu.au • 0450445373 • <https://telekhoa1812.github.io/My-Profile/> • <https://www.linkedin.com/in/dang-khoa-le-96a6332a8>

EDUCATION

Swinburne University of Technology

Software Engineer

SKILLS

Vietnamese, English, Mandarin

PROJECTS

Triage Response: LLM-Driven Solution and Drone Medication Delivery

- * Developed an AI-powered emergency triage system, leveraging LLMs (Whisper, Gemini, Qwen-VL) and trained on MIMIC-IV data for real-time clinical decision support.
- * Engineered integration of voice-triggered alerts with secure, rapid drone medication delivery, enhancing emergency response efficiency.

SkyLedge - Capstone Project

- * Spearheaded SkyLedge, an OBD-II predictive maintenance system, optimizing fleet servicing costs by 15% through real-time data and advanced fault prediction models. *
- Led a 4-person team, architecting the full-stack solution from hardware integration and cloud pipelines to a user-facing monitoring dashboard.

AI Tutoring

- * Architected and developed a full-stack AI tutoring web application, integrating multi-RAG LLMs (Gemini, Qwen) and custom KB embeddings for personalized, curriculum-aligned instruction. *
- Engineered a microservices platform using Next.js, FastAPI, and MongoDB, powering features such as level-based tutoring, automated grading, dynamic timetables, and content crawling.

Medical Chatbot

Engineered a cutting-edge AI medical chatbot leveraging Google Gemini Pro 2.5 for advanced diagnostic support and treatment recommendations, achieving 92% accuracy

- * in real-time query responses. Implemented a robust Retrieval Augmented Generation (RAG) architecture using NodeRAG over a multi-terabyte medical knowledge base, integrating FAISS (with GridFS) for ultra-low latency vector search and document retrieval, ensuring sub-200ms response times*.

Designed a sophisticated data management layer utilizing MongoDB for scalable storage, incorporating

dynamic Short-Term Memory (STM) and Long-Term Memory (LTM) history caching to deliver personalized user interactions and reduce redundant API calls by 30%*. Optimized model performance and efficiency through targeted fine-tuning of Gemini Pro on specialized medical datasets and applying post-training quantization, resulting in a 25% reduction in inference latency and 20% lower computational resource consumption*. Developed a high-performance FastAPI backend with integrated Natural Language Processing (NLP) for seamless multi-language support, coupled with a responsive Node.js frontend, enabling scalable and secure real-time communication via a RESTful API*.

EXPERIENCE

CPG Seedcom — Data Science | Team Leader June 2025 - September 2025

* Led a 6-person Data Science team to design and deploy a Vietnamese-language AI assistant, leveraging Gemini/Qwen3, Chain-of-Thought, and Agentic RAG to automate SQL query generation from complex business questions across MySQL and Oracle. * Achieved an 80% reduction in manual workforce by enabling autonomous analytics for sales, supply chain, and marketing through a multi-modal self-QA and Reinforcement Learning system, providing direct database-to-decision insights.

Pythera — AI Engineer | Part-time June 2025 - Now

* Led development of an agentic RAG module, integrating Electronic Medical Records (EMR) for Clara, a nationally recognized medical LLM+VLM assistant, critically enhancing diagnostic precision and medication safety. * Optimized core model training and benchmarking processes, while developing an intelligent chat memory retrieval system for robust contextual continuity in user interactions.

Nam Á Bank — Software Engineering - Data Science | Intern September 2024 - March 2025

* Engineered and deployed AI-driven solutions (e.g., OCR, LLM-based chatbot) using Vert.x and Java Android, advancing fintech digitalization and enhancing operational efficiency. * Designed and proposed a passport OCR system, incorporating synthetic data generation, estimated to reduce foreign customer onboarding time by X% through automated data extraction. * Advised on strategic product enhancements, including expanding international remittance services and redesigning mobile app interfaces for improved accessibility.