FDUCATION

SWINBURNE UNIVERSITY OF TECHNOLOGY

Bachelor of Software Engineer (with Honours)

Swinburne, Melbourne

2022-now

Relevant Coursework: Software Development, ML/DL, Web Full-stack Development, Mobile Application Development, Data Structures, Computer Vision, Data and Financial Analysis, Cybersecurity Deployment, Object Oriented Programming, Software Architecture, Cloud Architecture, Hardware Embedding.

SKILLS

Programming Languages: Python, Java, Kotlin, SQL, C++, C#, JavaScript, TypeScript.

Communication Languages: English, Vietnamese, Mandarin

Frameworks: FastAPI, Flask, Node.js, SpringBoot, Vue, React, Angular, Next.js.

AI/ML: TensorFlow, PyTorch, YOLO, ResNet, CNN, LLM, VLM, NLP, RAG architectures, Knowledge Distillation, CV, ASR, Machine Learning, Deep Learning, Synthesis, Data Analytics.

Tools: Docker, Kubernetes, MongoDB, MySQL, AWS, Azure, Google Cloud Console, Oracle, Jenkin, Jira, Git, Firebase.

EXPERIENCES

CPG Seedcom Data Science | Team Leader Ho Chi Minh City, Vietnam

June 2025 - September 2025

Leading a team of 6 to build a Vietnamese-language Al assistant, converts business questions into SQL queries across MySQL and Oracle, fusing multi-modal self-QA for Reinforcement Learning over failures and successes. Built with FastAPI, Gemini & Qwen3, MongoDB vectorbase, leveraging Chain-of-Thought and Agentic RAG for autonomous analytics in sales, supply chain, and marketing. The system reduces manual workforces by up to 80%, delivers instant, explainable insights, a unique plugin solution enables direct database-to-decision workflows without human intervention.

Pvthera Ho Chi Minh City, Vietnam

Al Engineer | Part-time

June 2025 - Now

Contributed to Clara, Vietnam nationally recognized medical LLM+VLM assistant trained on 400k+ doctor-annotated cases, including 210k+ X-rays, and synthesis. Led development of an agentic RAG module using EMR to improve diagnosis accuracy and medication safety. Also supported model training, benchmarking, and built intelligent chat memory retrieval for contextual continuity.

PopTech Ho Chi Minh City, Vietnam

ML/DL Engineer | Team Leader

Led a 4-member team to architect and build a scalable electricity forecasting and control application from IoT device. Integrated ML/DL (TFT, XGBoost), Reinforcement Learning (DQN), and LLMdriven advisory to predict billing, optimize device usage, and automate energy-saving decisions. Enabled multimodal interaction via ASR, text, and voice synthesis for a fully interactive MCP experience.

Software Engineering - Data Science | Intern

Ho Chi Minh City. Vietnam

December 2024 - March 2025

- Participate in fintech digitalization projects with full-stack development (Vert.x Backend, Java Android) and Al integration (CV, OCR, LLM, Chatbot). Proposed strategic enhancements: - Explore the high potential of remittance services targeting 5M+ overseas Vietnamese. Taking role in participation at Australian market, research for deployment.
- Redesign mobile app concept for a faster access, shifting complex tasks to web-services, deliver a user-friendly UI for accessibility-focused users. - Implement passport OCR to enable foreign customer onboarding with banking programs, overcome traditional limitation for identification with domestic ID. Implement synthetic data generation.

The Auburn Coffee Website Developer | Part-time Hawthorn East, Victoria, Australia April 2024 - Now

Developed the official website for a local commuter café at Auburn Station. Features include interactive menu, real-time pre-ordering, loyalty program with personalize QR code, and staff dashboard for order management, rating and feedback, sale analysis. Built with VueJS, Node.js (Vercel), database storage on MongoDB, and Firebase. Revolutionizing the business with digitalisation procedure, skyrocketing sale on an instance by 30%.

Ho Chi Minh City. Vietnam

Software Engineering - Data Science - Cloud Solution | Intern

December 2024 - March 2025

Analysing data, document, article for Sentiment Analysis and Business Overview with LLM models, integration with cloud computing architecture. Leading current team to leverage Obsei framework and Transformers for sentiment analysis, fact-checking, summarization, deliver actionable business insights for the clients.

VietDan JSC

Ho Chi Minh City, Vietnam

Software Engineering | Part-time

December 2024 - February 2025

Develop Water Meter OCR, a real-time water meter scanning solution using Next.js, Node.js, and Flask API, for government client. Self-mining and annotation of data, employs YOLOv5xu for detection and CV+OCR for text, hosting on Jetson Nano and AWS EC2. Production met the planned KPI to reduce manual labour by approximately 35%.

Hardware Technician | Part-time

September 2019 – December 2019

Hardware/Mechatronic technician specializing in motherboard replacement, repair and maintenance services. Have an in-hand experience with routers motherboard.

ALTA MEDIA

HCM City, Vietnam

Software Engineering | Intern Dev-Ops Engineer, deploying software architecture with Docker. Implement backend infrastructure for core software systems. December 2022 - February 2023

LEADERSHIPS

Triage Response: LLM-Driven Solution and Drone Medication Delivery

Developed and deployed an Al-powered emergency triage system integrating React Native (mobile), React (web portal), FastAPI (backend), Whisper-v3, Gemini 2.5, Qwen-VL, FAISS, and MongoDB with MIMIC-IV triage-response data to enable real-time voice-triggered alerts, clinical decision support, and secured-fast drone (Great Shark 330 PRO VTOL) medication delivery.

Sall-e: Autonomous Robot for Oceanic Garbage Collection

Led team of 5 engineers on software development and AI integration for Autonomous Robot combining computer vision, robotic techs, and navigation system. Directed cross-disciplinary R&D to deliver a scalable, Al-driven solution for ocean waste detection and collection. Data mining and self-train CNN and Transformer model, real-time access.

SkyLedge - Capstone Project

Hawthorn, Victoria, Australia

December 2024 - March 2025

Team Leader - AI and Backend Development

Led a team of 4 to build an OBD-II-based predictive maintenance system for SkyLedge, targeting reduced fleet servicing costs. Implemented real-time data logging with Raspberry Pi, automated cloud data engineering (FastAPI, Docker), and used K-Means clustering for driver behaviour analysis. Prototyped fault prediction with XGBoost, Random Forest, and LSTM, live monitoring dashboard.

PROJECTS

Al Tutoring

Sole founder-developer of an AI tutoring web app in active development. The platform combines LLMs (Gemini & Qwen in multi-RAG agentic architecture), custom KB embeddings from PDF and online sources crawler agent, deliver personalized, curriculum-aligned instruction. Features include level-based and virtual tutoring, automated grading, dynamic timetable roadmap with topic devision and covered materials, support website, video, and quizzes crawling from online sources. Built with React.js (Next.js, Tailwind, shaden/ui), FastAPI, MongoDB (GridFS), in microservices architecture.

Medical Chatbot

Developed an optimized AI medical chatbot with Gemini Pro 2.5, FAISS (with GridFS), NodeRAG over diagnosis & treatment medical dataset, real-time query. Data storage on MongoDB, leverage dynamic STM/LTM history caching. Incorporating model fine-tuning and quantization for efficiency. Deployed FastAPI backend with NLPs for multi-language support and NodeJS frontend, via RestAPI.

Golf Scorecard Scanner

A byproduct of VHandicap, with YOLOv8/YOLOv11 classification, VietOCR, PaddleOCR, MNIST CTC loss model for handwritten recognition. Integrates Qwen2.5-VL VLM agent for advanced reasoning.

Android app for vocabulary learning features mini-games, progress tracking and achievement, with real-time multi-language translation using Google ML Kit.

Roster Management System

Android application streamlining staff rostering and payment management with Kotlin, Node.js, and MariaDB. Deploying at Melbourne warehouse, minimise human-error and accounting labour.