

Huy Le

825-419-8287 | huyyylegia@gmail.com | linkedin.com/in/huylegia | github.com/giahuyyle

EDUCATION

University of Alberta

Bachelor of Science in Computer Science, Minor in Mathematics

Edmonton, AB

Expected May 2027

Coursework: Data Structures & Algorithms, Software Engineering, Operating Systems, Linear Algebra, Calculus

Awards: Dean's Honor Roll (2023–2024)

TECHNICAL SKILLS

Languages: Python, TypeScript/JavaScript, Java, SQL, C/C++/C#, HTML/CSS, Assembly

Backend/Testing: FastAPI, Flask, Node.js, REST APIs, ChromaDB, FAISS, WebSockets, U/I Testing, pytest, JUnit

Cloud/DevOps: Docker, AWS, Google Cloud Platform, Firebase (Auth/Firestore)

Frameworks/Libraries: React (Vite), Next.js, TailwindCSS, LangChain/LangGraph, pandas, NumPy, Matplotlib, PyTorch, scikit-learn, SQLAlchemy

EXPERIENCE

Applied AI/ML Software Engineer Intern

May 2025 – August 2025

FPT Software

Hanoi, VN

- Designed and implemented a lightweight, streamable **HTTP MCP** architecture for **OpenAI RAG** integration, reducing response latency by **35%** and enabling modular deployment.
- Profiled and optimized an end-to-end vector retrieval pipeline (LangChain + ChromaDB), improving query throughput by **42%** and overall chatbot responsiveness.
- Evaluated alternative retrieval approaches and proposed a **graph-based retrieval model**, boosting answer precision by **18%** over standard RAG baselines.

Software Engineer

June 2024 – August 2024

CodeDream Education

Hanoi, VN

- Translated stakeholder needs into technical requirements and delivered a **full-stack admin dashboard** (React, TypeScript, FastAPI), streamlining scheduling, tracking, and resource management for **4+ centers** and **50+ classes**.
- Performed preventative/corrective maintenance on a legacy PHP-based online judge (CDOJ), debugging core workflows and optimizing bottlenecks to **improve stability/performance by 14%** for **400+ students**.
- Led research and fine-tuning of **CodeLlama-based** models to detect AI-generated code, correctly identifying **55%** of AI-based submissions to support platform integrity.

Software Engineer Intern

May 2023 – September 2023

DIBIZ - Digital Business

Hanoi, VN

- Diagnosed bottlenecks and optimized data-handling algorithms for **RabbitMQ** streams (**500,000** customers/day), cutting processing time by **15%** and improving reliability of daily jobs.
- Implemented and integrated **10+ REST APIs** supporting user preference tracking and downstream personalization workflows, improving promotion relevance and engagement.

PROJECTS

TalkFlow | Next.js, TypeScript, Python, Flask, FasterWhisper, FAISS

June 2025 – August 2025

- Engineered a **full-stack** meeting platform with real-time messaging and audio/video using **WebSockets**, enabling low-latency collaboration across distributed users.
- Streamed live audio to backend and implemented a custom **speech-to-text pipeline** in **Python** (FasterWhisper), delivering real-time transcription.
- Integrated **FAISS** vector storage for conversation memory, achieving **less than 15ms** retrieval and generating end-of-meeting summaries.

ArenaIQ | React, JavaScript, Python, Flask, scikit-learn

April 2025 – June 2025

- Built an end-to-end **full-stack sports analytics dashboard** (React + Flask), translating user needs into features for live NBA tracking and Premier League match prediction views.
- Developed Python ML pipeline in **scikit-learn** (Decision Trees, Random Forests), including feature engineering and model evaluation, achieving **77% prediction accuracy** for EPL outcomes.
- Designed the data layer using **PostgreSQL** (historical) and **Firebase** (real-time), supporting reliable reads/writes across live match streams and **20 years** of player/team data.