



VO PHU THINH

AI ENGINEER

Address: Ho Chi Minh
Phone: 0868639882
Email: vophuthinhcm@gmail.com
Linkedin: [linkedin.com/in/vophuthinh](https://www.linkedin.com/in/vophuthinh)
GitHub: github.com/vophuthinh

EDUCATION

HCM UNIVERSITY OF TECHNOLOGY (HUTECH) Sep 2021 - Nov 2025

Bachelor of Information Technology

- Ho Chi Minh City University of Technology (HUTECH) & Open University Malaysia (OUM) Bachelor of Information Technology

UNIVERSITI MALAYSIA SARAWAK (UNIMAS) July 2023 – Aug 2023

Intensive English Language Mobility Programme

IRONHACK VIETNAM Mar 2022 - Aug 2023

Full Stack Web Developer

WORK EXPERIENCE

HPT VIETNAM CORPORATION Nov 2024 - Present

AI Engineer

- **Architected the Award-Winning "Lyly" (Executive AI Assistant):** Designed a multimodal system for real-time voice and text interaction, recognized with prestigious awards for technical innovation and low-latency performance.
- **Engineered End-to-End Vector Pipelines:** Developed automated workflows for strategic chunking and embedding, significantly enhancing semantic retrieval for enterprise knowledge management systems.
- **Orchestrated the HPT D-Day 2025 RAG Chatbot:** Deployed a high-accuracy RAG system providing real-time, consistent answers for event agenda, venue, registration, and speakers based on official data sources.
- **Optimized LLM Reliability in Microsoft Copilot:** Implemented advanced prompt engineering and grounding strategies within Microsoft Copilot to effectively mitigate hallucinations and ensure reliable outputs for business-critical workflows.
- **Co-developed the Award-Winning HIH (HotNews Intelligence Hub):** Built an automated intelligence system featuring "Sóc Dò La" for news vectorization and summarization, earning top honors for its practical impact on executive decision-making.

PROJECTS

LYLY - AI ASSISTANT FOR EXECUTIVE BOARD Dec 2025 - Jan 2026

Number of members: 04 (AI Engineer)

Description: Developed an enterprise AI chatbot named Lyly with real-time voice and text interaction to support internal knowledge management and business operations at HPT Corporation.

Feature:

- Implemented a voice-first AI assistant with bidirectional speech-to-speech communication using WebSocket for low-latency interaction.
- Built a text chat system with real-time streaming responses (SSE) and support for rich content such as charts, tables, and dashboards.
- Integrated a long-term memory system using vector database for context-aware and personalized responses.
- Developed a screen sharing feature allowing the AI to understand visual context and assist users based on on-screen content.
- Applied data visualization tools to generate dynamic charts and reports from enterprise data.
- Implemented security mechanisms including API key authentication, rate limiting, and PII detection to ensure enterprise-level data protection.

Technology:

- Front-end: Next.js 14, React 18, TypeScript, Tailwind CSS, Framer Motion.
- Back-end: Python, FastAPI, WebSocket, SSE (Server-Sent Events).
- AI & Data: Gemini 2.5 Native Audio, google-genai SDK, Qdrant Vector Database, Mem0 Memory.
- Database: PostgreSQL.
- Tools & Integration: MCP Toolbox for Databases, Sema4 MCP.
- Testing: Vitest (Frontend), Pytest (Backend), React Testing Library.
- Engineering: Git/GitHub, VS Code, Docker Compose.

Number of members: 01 (Final Year Project).

Description: Developed a full-stack e-commerce web application with integrated sentiment analysis and real-time messaging to enhance user experience and business management.

Feature:

- Implemented **user management** with secure registration, authentication, and role-based access (customer, seller, admin).
- Built **e-commerce functionalities** including product listing, search, secure checkout, and order tracking.
- Applied **sentiment analysis** to classify customer reviews into Positive, Neutral, and Negative for business insights.
- Developed **order and transaction management** allowing customers to track orders and sellers to manage order statuses and withdrawal requests.
- Integrated a **real-time messaging system** using WebSocket for seamless communication between customers and sellers.

Technology:

- Front-end: ReactJS, JavaScript (ES6), Tailwind CSS.
- Back-end: NodeJS, Express, WebSocket.
- Database: MySQL, Sequence.
- API & Integration: RESTful API, Payment Gateway Integration.
- Engineering: GitHub, Postman v11, XAMPP.

SKILLS

- **Artificial Intelligence & NLP:** LLMs, RAG, Prompt Engineering, Grounded Answering (Hallucination Reduction), Embeddings, Semantic Search, Conversational AI (Text & Voice), ASR/TTS.
- **ML Frameworks & Libraries:** PyTorch, scikit-learn, Hugging Face Transformers, Sentence-Transformers.
- **Vector DB & Automation:** Qdrant, n8n, Node-RED, UIPath, Airflow.
- **Backend & APIs:** Python, SQL, FastAPI, NodeJS, Express, RESTful API, WebSocket, SSE.
- **Enterprise Integration:** Microsoft Graph API, SharePoint Integration.
- **Frontend:** ReactJS, Tailwind CSS, JavaScript (ES6), Responsive UI, AJAX, JSON.
- **Databases:** Microsoft SQL Server, MySQL, Oracle, MongoDB.
- **DevOps & Version Control:** Docker, Git, Bitbucket.
- **Tools:** Jira, Trello, Postman, Figma, draw.io, Knack.
- **Languages:** English.

AWARDS & CERTIFICATIONS

NVIDIA DEEP LEARNING INSTITUTE

Oct 2025 - Dec 2025

- Certificate of Competency: Building LLM Applications With Prompt Engineering
- Certificate of Competency: Getting Started with Deep Learning

IBM - COURSERA

Sep 2025

- Professional Certificate: Machine Learning with Python

HPT AI IN THE WORKPLACE COMPETITION

Jan 2026

- **Second Prize:** Project "AI Agent hỗ trợ điều hành, quản trị" (Lyly Assistant)
- **Third Prize:** Project "H.I.H" (HotNews Intelligence Hub)