Tran Minh Duong

AI Engineer Intern

Hanoi, Vietnam   minhduongqo@gmail.com   0961 264 447   tmdeptrai.vercel.app

tran-minh-duong   tmdeptrai

About Me

I design, train, and deploy ML systems with a focus on reliability, monitoring, and real-world constraints. Experienced in RAG pipelines, LLM fine-tuning, model evaluation, containerized inference and cloud deployment.

Looking for AI Engineering / MLOps roles where I can ship and improve model delivery.

Education

BSc   **La Rochelle University**, Computer Science   La Rochelle, France
- Selected for International Exchange Program   Sept 2025 – present

BSc   **University of Science and Technology of Hanoi**, Information and Communication Technology (Double Degree)   Hanoi, Vietnam   Sept 2023 – July 2025
- GPA: 18.48/20 (≈ 3.89/4), ranked 1st in double degree cohort

Experience

ICTLab - USTH, Research Assistant   Hanoi, Vietnam   Mar 2025 – July 2025

Contributed to research on **context-aware fire detection** using Visual Language Models
- Published the fire context awareness dataset ⧉ to benchmark multimodal VLMs
- Evaluated Qwen2.5VL, InternVL3 and Gemma3 by implementing pipelines to compare various metrics and inference latency
- Fine-tuned VLMs using **PEFT/LoRA** to boost model accuracy up to 85% - 90%

Projects

Legal Contract Analyzer ⧉   July 2025 – Aug 2025

RAG Chatbot for understanding and extracting key information from legal documents
- Designed and implemented end-to-end MLOps pipeline with CI/CD and monitoring
- Uses RAG with semantic search, embeddings + LLMs to retrieve relevant sections
- **Techstack**: Github Actions, FastAPI, vLLM, ChromaDB, Docker, Prometheus/Grafana

Fire Detection System ⧉   Mar 2025 – Apr 2025

(Group Project Leader) IoT-based real-time fire detection and alerting system.
- YOLOv8 trained on 1k+ annotated flame/smoke images (0.85 mAP0.5 and Precision = 0.95)
- Integrated Dockerized inference with Telegram alerts for live notifications
- **Techstack**: Roboflow, Ultralytics YOLOv8, Docker, OpenCV

Skills

Programming Languages: Python (proficient), Java, C++, Matlab, Bash

Framework / Libraries: PyTorch, Hugging Face transformers, FastAPI, Langchain

Tools: Git, Docker, CI/CD, Prometheus, Grafana, AWS

Languages: English C1 (IELTS 7.5), French B2 (DELF B2)

Certifications

AWS Certified Solution Architect Associate ⧉