NGO GIA PHONG

Ha Noi, Viet Nam

OVERVIEW

- Full-stack Web Developer with strong skills in AI integration and LLM-based system development.
- Fast learning, self-motivated.
- Strong background in maths and statistics.

SKILLS

- **OS :** Ubuntu, MacOS
- **Programming Languages:** HTML/CSS, Javascript, Typescript, Python, C++, Bash Shell
- **Web Technologies / Frameworks:** Reactjs, Nextjs, MUI, TailwindCSS, FastAPI, Flask
- **Tools & Technologies:** AWS, Docker, Gitlab CI/CD, VScode, Figma
- **Databases :** Postgres, MySQL, Redis, Faiss, Qdrant

WORK EXPERIENCE

Software Engineer   04/2023 - now
- Develop and maintain web application systems using modern JavaScript frameworks across both frontend and backend, with LLM integration for building AI-driven systems.

PROJECT

Rabiloo Co., Ltd | *Software Engineer*   Hanoi | 04.2023-Now

TestBot – Automated Test Generator

Position: Fullstack developer   Aug 2025 - Now
Description:
- System that automatically generates test cases and test scripts from uploaded documents with CMS for management.

My responsibility:
- Research, build and integration AI model

- Design full pipeline for document processing and generation
- Build FE/BE and integrate LLM-based generation flows
- Implement CMS for test case & test script management
- Build queuing system to reduce latency and improve UX
- Fix bugs and implement change requests

Team size: 6
Technologies: LLM, LangChain, Python, FastAPI, PostgreSQL, Docker, AWS

Automatic Ad Generation System

Position: Researcher and Software Engineer   May 2025 - Aug 2025
Description:
- AI-powered chatbot system that generates advertising content automatically from user descriptions and outputs structured formats compatible with Google Ads.

My responsibility:
- Research, build and integration AI model
- Design system
- Build LLM-based generation pipelines
- Deploy services on AWS
- Fix bugs and implement change requests

Team size: 4
Technologies: LLM, LangChain, AWS, Python, FastAPI, Docker, vLLM

AI-Driven Document Management & Retrieval Platform

Position: Fullstack developer   Nov 2024 - Apr 2025
Description:
- A platform that manages project documents with role-based access control, approval workflows, version tracking, and an integrated RAG system for querying uploaded files through natural language.

My responsibility:
- Develop document management features: user/role management, document creation, approval flow, version tracking
- Design chatbot pipelines for structured and unstructured data
- Build chatbot and retrieval workflows
- Deploy and optimize model serving on cloud
- Work with customers to clarify requirements
- Fix bugs and implement change requests

Team size: 8
Technologies: RAG, LLM, LangChain, Transformers, PyTorch, Faiss, PostgreSQL, NextJS, AWS, vLLM, Docker

Translation System

Position: Fullstack developer   Oct 2024 - Mar 2025
Description:
- A translation platform that supports text, file, image, and voice translation with customizable translation settings.

My responsibility:

- Manage users, roles, groups
- Integrate with translation service and other services
- Allow user translate text, file, image, voice
- Allow user create setting to translate such as: domain, dictionary, model, etc

Team size: 6

Technologies: ReactJS, NextJS, MUI, MongoDB

Healthcare Information Extraction

Position: Researcher and Software Engineer   June 2024 - Oct 2024

Description:
- Extract important information from French users' healthcare documents (image, pdf, doc, handwriting)
- The tasks performed: Object Detection, Text Detection, Text Recognition, Table Recognition, Information Extraction

My responsibility:
- Search and collect data sources
- Synthetics dataset
- Research, build and integration AI model
- Design system
- Serving LLM models
- Fix bugs and implement change requests

Team size: 6

Technologies: LLM fine tuning, YOLO, transformer, Pytorch, PaddlePaddle, Docker, Conda, Minio, Python, FastAPI, Onnx, OpenCV

Building Owner & Vendor Matching System

Position: FE developer   May 2024 - Aug 2024

Description:
- A system connecting building owners and vendors through building projects and enabling interaction via chat.

My responsibility:
- Manage users, roles
- Allow building owner create building, project of building
- Allow vendor view and join public project of building
- Allow building owner and vendor can interact each other via chat

Team size: 4

Technologies: NextJS, MUI, TailwindCSS

Multilingual Document OCR System

Position: Researcher and Software Engineer   Nov 2023 - May 2024

Description:
- The system recognizes document information from images
- The tasks performed: Text Detection, Text Recognition, Information Extraction

My responsibility:
- Search and collect data sources
- Research, build and integration AI model
- Design system

- Fix bugs and implement change requests

Team size: 4

Technologies: LLM, Transformers, PaddleOCR, ONNX, Docker

Digital AI Signage

Position: Researcher and Software Engineer   Apr 2023 - Nov 2023

Description

- Analyze the attributes of age, gender, and emotions of shoppers when viewing ads
- The tasks performed: Face Detection, Face Recognition, Face Age, Face Attribute, Face Quality, model quantization, model pruning

My responsibility

- Search and collect data sources
- Research, build and integration AI model
- Design system
- Fix bugs and implement change requests

Team size: 6

Technologies: Transformer, CNN, SSD, YOLO, Docker, Conda, MySQL, Minio, Python, FastAPI, Onnx, OpenCV, C++, ncnn

EKYC

Position: Researcher and Software Engineer   Apr 2023 – Aug 2023

Description

- Build integrated attribute modules on mobile applications to orient warnings and guide users.
- Automatic personal information extraction system from identification documents.
- Biometric-based identity verification
- The tasks performed: The tasks performed: Face Detection, Face Recognition, Face Anti-Spoofing, Card Detection, Corner Detection, Card Quality Checking, Text Detection, Text Recognition, Information Extraction, Face Quality Checking.

My responsibility

- Search and collect data sources
- Research, build, and integrate AI model
- Fix bugs and implement change requests

Team size: 4

Technologies: Pytorch, OpenCV, Onnx, FastAPI, LMDB, Docker, Paddle, U2Net, yolov5, Yolov8

T-com Technology Company | *AI Intern*   Hanoi | 01.2023-04.2023

Position: AI Intern

Description:

Developed and evaluated Face Recognition models using company employee datasets, enhanced dataset quality through cleaning and merging, and conducted research on object tracking problems.

My Responsibilities:

Trained and evaluated Face Recognition models on proprietary employee datasets.

Cleaned and merged large-scale face datasets to improve recognition accuracy.

Researched solutions for object tracking challenges.

Technologies: Python, OpenCV, Image Processing, PyTorch

EDUCATION - QUALIFICATIONS

Ha Noi University of Science and Technology   **2018 - 2023**

- Toeic: 700/990