PhongMT

📍 Hanoi   ✉ phong.mt184172@gmail.com   📞 +84 396651152   🔗 https://maiphong0411.github.io/phongkamy/

○ maiphong0411

Introduction

I am Peter, Data Scientist with expertise in AI, NLP, and marketing analytics. Experienced in analyzing customer interactions at Johnson & Johnson to drive data-driven strategies and then turning complex data into actionable insights to support business growth. Skilled in building **GenAI solutions**, including Retrieval-Augmented Generation (RAG) systems and fine-tuning large language models (LLMs). Seeking new opportunities to apply my skills in impactful projects.

Education

EG   **Hanoi University Science and Technology**, Information Technology   Sept 2018 – Oct 2023

- GPA: 3.0/4.0
- **Coursework:** Data Structures and Algorithms, Operating System , Object-Oriented Languages and Theory, Artificial Intelligence, Knowledge Engineering, Structured Programming , Database

Vingroup, AI Trainee Program   Jul 2023 - May 2024

- **Coursework:** Python for AI, Machine learning, Deep learning, Computer Vision, Natural Language Processing, Machine learning operation, AI Ethics, Data Engineering.

Experience

Johnson and Johnson (AU), Data Scientist   Feb 2025 - Present

- Design and build automatic content tagging system for Japan, Australia and many regions. The system contributed to measurement of content materials and insights generation.
- Ad-hoc analysis: find valuable insights under data, the tactic strategies for marketing team to launch campaign.
- PDF translation: translate PDF materials from source language to target language preserving format of the documents.

VMO Holdings, AI Engineer   May 2024 - Present

- Develop a chatbot system for the company that helps stakeholders and investors learn more about the company.
- Remediate the finance description system.
- Doing research and fine tuning Large Language Model (Llama 3, Mistral).
- Contribute to developing SQL agent for an enterprise utilizing LLM models for marketing team who has ability to retrieve data by natural language.
- Hosting some techinical sharings about MCP, RAG, Tokenizer, Prompt Engineering.

VinBigData - VinBrain, Applied Scientist   Jul 2023 - May 2024

- Develop streaming ASR system which was applied to the enterprise's application.
- Creating synthetic data for automatic speed recognition.

VEF Academy, Teaching Assistant   April 2022 - Oct 2023

- Instruct the students to do hands-on data analysis projects and assignments (EDA, Machine learning)
- Collaborate with the main teachers to design the lab and lessons

Projects

Graduation Thesis: Machine translation for multiple domains

- The thesis was instructed by Associate Prof Le Thanh Huong
- Develop a translation model (based on MBart) to convert English to Vietnamese language with low resources on multiple domains like: medical, technology
- Total score: 9.3 / 10

User Management Via Chat   user-management-via-chat ↗

- Apply the latest technology to help colleagues reduce time when creating new user in the server.
- Keyword: Model Context Protocol, Claude Desktop, Claude Model

Medical Chatbot

- Chatbot helps users clarify symptoms and answer questions about the use of drugs with common diseases.
- The project was awarded by VietAI. Link to reference: Medical Chatbot ↗

Mini RAG   miniRAG ↗

- Build an autonomous agent with RAG from scratch. The agent has the ability to make the questions to clarify the user's intent.

Code Generation   Code Generation ↗

- Train LLM with low resource for code generation
- Keyword: DDP, Lora, Mixed Precision Training

Certificates

VietAI: LLMs and Industry Practices 01

VietAI: Advanced Natural Language Processing

Datascienceworld.Kan: Machine learning operation

Technologies

Languages: Python, C, Java

Technologies: Langchain, Langgraph, Vector database, Model Context Protocol, LLM, Prompt Engineering, FastAPI, Docker, Git.

Languages

Vietnamese: Native proficiency

English: Intermediate (Currently I am working at global company)

Japanese: JLPT N3