

# Bùi Trần Công Minh

**Date of birth:** 09/01/2000

**Address:** Ho Chi Minh, Vietnam

**Phone number:** 0353494569

**Email address:** minh.btrc@gmail.com

**Linkedin:** <https://www.linkedin.com/in/minhbtrcm00>

**Github:** <https://github.com/minhbtrc>

## Profile

- I am an AI Engineer with extensive expertise in machine learning and deep learning, specializing in natural language processing (NLP) and Transformer-based architectures. My experience includes applying Large Language Models (LLMs) to diverse tasks such as question generation, sentiment analysis, and data extraction.
- I am passionate about optimizing LLM performance using advanced techniques and frameworks like Langchain. I am particularly eager to expand my skills in deploying LLMs to production environments and am enthusiastic about opportunities to learn and contribute in this area.

## Education

08/2018 – 04/2023  
Ho Chi Minh, Vietnam

**Computer Science | Bachelor**  
**Ho Chi Minh University of Technology**

08/2015 – 08/2018  
Gia Lai, Vietnam

**Mathematics**  
**Hung Vuong High School for the Gifted**

## Work experience

09/2024 – present  
Ho Chi Minh City,  
Vietnam

**AI Engineer**  
**FPT Software AI Center**

10/2023 – 11/2024  
Ho Chi Minh, Vietnam

**AI Engineer**  
**SPARTAN**

### PDF Data Mining Project

#### **PDF Parser Module:**

- Led the development of the PDF Parser module, researching and implementing advanced document parsing techniques.
- Enhanced parsing processes by integrating Python-based libraries and leveraging Azure Document Intelligence for improved accuracy and efficiency.

#### **LLM Integration for Data Mining:**

- Leveraged Large Language Models (LLMs) to extract and format data from PDF files, producing structured outputs.
- Collaborated with the team to design and implement a service that schedules API calls to LLM Cloud services, optimizing the service to avoid hitting rate limits.

#### **Full Pipeline Development & Deployment:**

- Worked closely with the team to create a comprehensive pipeline for the service, ensuring seamless integration and functionality.
- Contributed to the preparation of Kubernetes (K8s) deployment scripts, assisting the team in deploying services and applications into Google Cloud Platform (GCP).

## Work experience

---

07/2021 – 10/2023  
Ho Chi Minh, Vietnam

### Backend Development:

- Led the development of an application that allows users to drag and drop nodes to create a flow for handling PDF parsing, leveraging the Langflow open-source project.
- Implemented, handled, and improved APIs for the application, ensuring robust and scalable backend services.

### AI Engineer

#### TRUONG MINH THINH TECHNOLOGY JOINT STOCK COMPANY

### Scenario Design & Implementation for Chatbot Team:

- Designed and implemented chatbot scenarios based on pre-defined scripts, ensuring seamless user interaction.
- Maintained and optimized chatbot services, guaranteeing smooth and logical system operations.

### Advanced AI Techniques & Model Development:

- Researched and re-implemented state-of-the-art techniques for specific AI tasks, including:
- Question Generation: Developed models using BARTPho and Marian architectures, applying the pointer technique to handle out-of-vocabulary tokens during response generation.
- Sentiment Analysis: Built and fine-tuned sentiment analysis models to classify user messages, enabling the visualization of chatbot response effectiveness and usefulness.
- Experienced in fine-tuning LLM pretrained models using techniques such as Hugging Face's SFT Trainer and LoRA, focusing on hyperparameter optimization for enhanced model performance.

### Performance Optimization:

- Utilized ONNX and TorchScript to optimize AI models, significantly improving inference time and overall system performance.
- Managed and versioned data using DVC, ensuring robust and reproducible model training pipelines.

### Prompt Engineering:

- Applied prompt engineering techniques to refine model outputs, enhancing the chatbot's ability to understand and respond to user inputs.

### Langchain Integration:

- Leveraged Langchain to handle scenarios outside of predefined flows, enhancing the user experience by providing more flexible and dynamic chatbot interactions.

## Projects

---

Ho Chi Minh, Vietnam

### Lead Covid Chatbot

- Spearheaded the development of a consultancy chatbot aimed at assisting SARS-COVID-2 patients with symptom assessment and self-care advice at home
- Provided essential Covid-19 information and diagnosed potential infections based on user-provided data
- Translated an English dataset into Vietnamese using Selenium to ensure the chatbot's accessibility and effectiveness for Vietnamese users
- Implemented machine learning models (KNN, SVM) for the Intent

## Projects

---

Classification module to enhance the chatbot's ability to accurately understand user queries

- Developed a Named Entity Recognition (NER) system using PhoBERT and Conditional Random Fields (CRF) to extract pertinent entities from user input.

Ho Chi Minh, Vietnam

### Lead Langchain Chatbot

#### Project Overview:

- Developed a chatbot using the Langchain framework, integrated with Vertex AI or OpenAI API for the base LLM.
- Implemented MongoDB for memory management and utilized Gradio and Langchain UI for the user interface, supporting LangChain's streaming capabilities.
- Enabled prompt versioning and tracing through LangSmith, and allowed user customization of the bot's personality.
- Incorporated Microsoft Presidio for data anonymization, ensuring protection of personally identifiable information (PII) before interactions with the LLM.

**Project Link:** <https://github.com/minhbtrc/langchain-chatbot>

## Hobbies

---



Book



Paper research



Football

## Skills

---

### - PROGRAMING SKILLS

FastAPI



Python



Docker/K8s



### - NLP/LLM

Langchain



Prompting



Pytorch



## Certificates

---

08/2023

### LangChain & Vector Databases in Production ActiveLoop

<https://learn.activeloop.ai/certificates/snnl0bjoa4>

08/2023

### Generative AI with Large Language Models Coursera

<https://www.coursera.org/account/accomplishments/certificate/M8QJYLYC9FC>  
D

09/2023

### Generative AI Fundamentals Google Cloud Skills Boost

[https://www.cloudskillsboost.google/public\\_profiles/62daf634-c499-4c31-8ed2-bf1c0d6a9b15/badges/5406285?](https://www.cloudskillsboost.google/public_profiles/62daf634-c499-4c31-8ed2-bf1c0d6a9b15/badges/5406285?utm_medium=social&utm_source=linkedin&utm_campaign=ql-social-share)  
[utm\\_medium=social&utm\\_source=linkedin&utm\\_campaign=ql-social-share](https://www.cloudskillsboost.google/public_profiles/62daf634-c499-4c31-8ed2-bf1c0d6a9b15/badges/5406285?utm_medium=social&utm_source=linkedin&utm_campaign=ql-social-share)

03/2022

**Deep Neural Networks with PyTorch**  
**IBM**

[https://www.credly.com/badges/1a28ac75-fe07-46c8-af42-d6c655cde21a?source=linked\\_in\\_profile](https://www.credly.com/badges/1a28ac75-fe07-46c8-af42-d6c655cde21a?source=linked_in_profile)

- Completed the Deep Neural Networks with PyTorch certification from IBM, demonstrating proficiency in implementing and enhancing neural networks with the PyTorch framework.
- Acquired expertise in optimizing neural networks and utilizing PyTorch through hands-on training and projects.

03/2022

**Natural Language Processing Specialization**  
**Coursera**

<https://www.coursera.org/account/accomplishments/specialization/certificate/8JVX3M2H36T2>