NGUYEN DUONG THANH DU

+84 886 617 378 Github Linkedin Google Cloud Profile Email

EDUCATION

IUH - Industrial University of Ho Chi Minh City **10/2020 - 04/2025 (Expected)**
Bachelor Dergree in Computer Science Programs. *GPA: 3.62/4.00*

TECHNICAL SKILLS

Programming Languages: Python
Libraries and Tools: PyTorch, Tensorflow, Sklearn, Pandas, Numpy, OpenCV
ML Architectures: CNN, RNN (Regions), Fast-RCNN, Faster-RCNN, Transformers(BERT)
LLMs Libraries: LlamaIndex, LangChain, Guidance
Other: Git, Docker, FastAPI, AWS, Google Cloud (Vertex AI)

WORK EXPERIENCE

AI Engineer
BlueBolt Software, Ho Chi Minh City, VietNam *05/2023 - 04/2024*

* Built **dashboards** for analyzing company data, tracking performance, and providing evaluative insights to enhance efficiency and productivity. Created visually impactful reports to **highlight data insights** and improve data visualization capabilities.
* Deployed **Odoo Community (ERP)** system on **AWS**, managing related servers hosting Odoo on the AWS platform. Developed **automated scripts** for monthly backup of Odoo database. **Managed and maintained** operational processes of Odoo.
* Developed a **Retrieval Augmented Generation (RAG) Chatbot** using large language models to assist users with **medical health** inquiries by crawling and processing medical data from Vinmec, building a **vector database** stored in PostgreSQL (pgvector), and implementing a RAG pipeline with **ChatGPT 3.5** as the language model.
* **Enhanced the RAG pipeline** with advanced techniques such as re-ranking, chain of thought prompting, and hybrid search to improve the chatbot's performance and accuracy in retrieving and generating relevant medical information.
* **Deployed the entire RAG** Chatbot pipeline to production, enabling users to interact with the chatbot and receive assistance for their medical health-related questions and information needs.
* Configured a Jetson Nano embedded AI platform for **face detection** tasks, including loading and **optimizing pre-trained face detection** models onto the Jetson Nano's GPU for efficient **real-time** facial recognition and analysis.

PROJECTS

Medical Document Retrieval Chatbot **Demo**

* Built a Retrieval Augmented Generation (RAG) chatbot utilizing Llama-Index, PostgreSQL vector database, and GPT-3.5 for medical document retrieval.
* Implemented data embedding, indexing, and advanced text retrieval techniques like keyword search, semantic similarity, and context-awareness.
* Developed in Python 3.10 using Llama-Index, PostgreSQL vectors database, and OpenAI GPT-3.5.

K-means Parallel **Demo**

* Developed parallel K-means clustering using Map-Reduce on Hadoop with Python.
* Leveraged Hadoop for distributed computing and parallelization of the K-means clustering algorithm across a cluster.
* Provided data generation, implementation execution, and result comparison instructions.

ACHIEVEMENTS AND EXTRACURRICULAR ACTIVITIES

Term-based scholarship *at IUH - Industrial University of Ho Chi Minh City* *2020 - 2024*
Certificate *Google Data Analytics Specialization at Coursera* *7/2023*
QuanQuanGCP *for learning Google Cloud Skill* *3/2024*