

TAI THANH-LE

Thu Duc city, Ho Chi Minh city

Phone number: (+84) · 774 · 707 · 297 ◇ Email: taifa2907@gmail.com

[LinkedIn](#)|[Github](#)|[Portfolio](#)

EDUCATION

HCMC University of Technology and Education

June 2022-Present

B.S. in Embedded System and Internet of Things

Overall GPA: 8.1/10 (3.3/4.0)

IELTS Academic: 7.0

SKILLS

- **Programming Languages:** Python, SQL, C/C++, C#, JavaScript
- **Frameworks:** Scikit-learn, Tensorflow, Keras, PyTorch, Langchain, FAISS, MLflow, OpenCV, Hugging Face Transformers, Flask, Streamlit.
- **Technologies:** Git, Docker, Linux, LaTeX, AWS.
- **Soft Skills:** Analytical Thinking, Collaboration, Strong Communication, Problem Solving, Quick Adaptation, Rapid learner.

PROJECTS

Unsupervised Industrial Anomaly Detection with Vision Graph Neural Network-based Autoencoder | [Github](#) January 2025 - Present

Technologies: PyTorch, FAISS, MLflow, Scikit-learn

- Fine-tuned Vision GNN backbone for better features extraction and enhance the ability to capture spatial dependencies in image.
- Combine with KNN-based algorithms for identify anomalies at pixel-level, enhance the segmentation ability by **31.4%**.
- Utilized FAISS for normal features savings and having better retrieval abilities.
- Archived **0.98** in anomaly instance AUROC, **0.93** pixel-wise AUROC, **0.931** anomaly-pixel AUROC throughout 15 difference industrial product in a real-world dataset that is MVTec AD.

Text-to-Image with Stable Diffusion | [Github](#)

February 2025 - March 2025

Technologies: PyTorch, Diffusion Model, Hugging Face Transformers

- Successfully reimplemented the Stable Diffusion model from the original research paper "High-Resolution Image Synthesis with Latent Diffusion Models" using PyTorch.
- Built the model architecture from scratch, including the U-Net structure for denoising and integrating attention mechanisms for improved image quality.
- Deploy Created an intuitive web application using Streamlit that allows users to input text prompts for image generation, upload images for image-to-image generation with adjustable strength parameters and view generated images in real-time.

Advanced Driver-Assistance System | [Github](#)

September 2024 - December 2024

Technologies: PyTorch, Tensorflow, Keras, YOLO, OpenCV, Flask, Image Processing

- Fine-tuned YOLOv8 on self-collected dataset of traffic in Vietnam for detecting traffic signs and vehicles, archiving **0.79** mAP50, **0.76** precision and **0.77** recall.
- Designed and implemented in Tensorflow a lighter U-Net based model for lane segmentation, delivered **97.2%** pixel-wise accuracy.
- Formatted model to Tensorflowlite format combined with multi-threading technique to archive real-time performance which is **less than 30 ms/frame**.

Vietnamese Chatbot using Retrieval-Augmented Generation | Github May 2024 - June 2024
Technologies: Python, LLM, Langchain, RAG, FAISS, Hugging Face Transformers

- Conduct an intelligent conversational agent for documentation inquiries in Vietnamese.
- Utilized LangChain to manage interactions between large language models (LLMs) and retrieval systems for more accurate information fetching and coherent response generation.
- Used FAISS for efficient vector storage and retrieval of content got from PDF file user uploaded.
- Built a user-friendly interface using Streamlit for seamless interaction.

EXPERIENCE

Researcher Assistance

March 2025 - Present

University of Technology and Education

- Conducting comprehensive literature reviews on advanced diffusion models for conditional image generation, including **DPT-T2I, Instance Diffusion, Ranni and Readout Guidance**.
- Reimplementing state-of-the-art methodologies to ensure fidelity and performance in image generation tasks.
- Executing fine-tuning techniques to optimize model performance, resulting in significant improvements in image quality and generation efficiency.
- Collaborating with interdisciplinary teams to analyze experimental results, employing rigorous statistical methods to refine and enhance methodological frameworks.

Team Leader

October 2024 - December 2024

Practical Machine Learning and Artificial Intelligence Course Project

University of Technology and Education

Technologies: Python, PyTorch, FAISS, Scikit-learn

- Led a team of three to reimplement and optimize PatchCore model proposed in the paper "Towards Total Recall in Industrial Anomaly Detection".
- Trained the model from scratch for industrial anomaly detection task on MVTec AD dataset and a custom dataset of colorbured cans in a spinning wheel.
- Integrated FAISS for more efficient features retrieval, increase the inference speed **20 times** faster than using non-vector databases.
- Archived **0.973** in anomaly instance AUROC, **0.929** pixel-wise AUROC and **0.921** anomaly-pixel AUROC.

Team Leader

October 2023 - November 2023

Foundation of Data Science Course Project

University of Technology and Education

Technologies: Python, Scikit-learn, Pandas, NumPy, KMeans, Matplotlib

- Leading a team of eight members for clustering customer locations for setting up dairy products recycling collection points.
- Developing and hyperparameters tuning by performing Elbow and Silhouette Analysis for K-Means.
- Calculate and analysis cost and recycle rates, archived the cost of **\$419,216** (less **\$10,000** than other groups). Got a **93.72%** in total Recycle Circular Inflow. For recycle rates, carton containers(**1.2%** increment), glass bottles(**100%** increment), plastic bottle(**2.91%** increment).

CERTIFICATES

Data Analysis with Python

September-2023

Grade: 100%

IBM

- Mastered data preprocessing, perform exploratory data analysis using Pandas and NumPy, and build regression model with Scikit-learn.
- Developed skills in data manipulation, visualization, and predictive modeling.

Machine Learning with Python

Septemper-2023

Grade: 100%

IBM

- Gained basic knowledge of machine learning algorithms, including: Linear Regression, Logistic Regression, Decision Trees, Random Forests, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), K-Means Clustering.
- Developed skills in regression analysis and model evaluation using SciPy and scikit-learn.

Python Project for Data Science

Septemper-2023

Grade: 90%

IBM

- Conducted an EDA on a real-world project, applied Python fundamentals techniques.
- Engineered interactive dashboards using Pandas and Beautiful Soup within Jupyter Notebooks; extracted key data insights that identified the top three factors contributing to project delays, enhancing decision-making processes.

Python for Data Science, AI & Development

Septemper-2023

Grade: 100%

IBM

- Learned Python programming fundamentals, including variables, data structures, and functions. Learned Python programming fundamentals, including variables, datastructures, and functions.
- Developed proficiency in libraries like Pandas and NumPy, and utilized Jupyter Notebooks for coding.
- Gained skills in data access and web scraping using APIs

What is Data Science

Septemper-2023

Grade: 100%

IBM

- Explored the significance of data science in the modern world and various career paths within the field.
- Summarized insights from experienced professionals and discussed why data science is a leading job in the 21st century.