KHANG BUI TRAN DUY

AI ENGINEER | DATA SCIENCE

(+84) 815 594 665 | khang.buitranduycse@hcmut.edu.vn | khangbkk23 | khangbkk23

I. ABOUT ME

Summary: Third-year Computer Science student at HCMUT with a foundation in Machine Learning, Deep Learning, Computer Vision, Natural Language Processing, currently expanding into Reinforcement Learning.
Short-term goal: To obtain an AI Engineer position where I can collaborate with experienced professionals, contribute to real-world AI systems, and strengthen my technical and problem-solving skills.
Long-term goal: To become a core AI Engineer and pursue advanced international education. My ultimate aim is to acquire cutting-edge global knowledge and return to drive high-impact R&D initiatives for the company.

II. EDUCATION

Ho Chi Minh City University of Technology - Vietnam National University, HCMC *Sep 2023 - Apr 2027*

Bachelor of Engineering in Computer Science, Specializing in Applied Artificial Intelligence

CPA: **3.71/4.00**

III. RESEARCH EXPERIENCE

University of Technology - VNU-HCM - Faculty of Computer Science and Engineering *May 2025 - Present*

Research Team Member -Topic: Optimization of Continual Learning Strategy for RL and DNN *Repository*

Subject: Nested Learning & Continuum Memory System

- **Architected ViT-CMS**: Integrated **Nested Learning** blocks with a **Continuum Memory System** on a ViT-Base backbone to **mitigate catastrophic forgetting** without storing raw data.
- **Nested Optimizer:** Designed a **multi-frequency update schedule** and **Temporal isolation** mechanism to ensure precise hierarchical weight organization.
- **Performance:** Achieved **73.1% F1-score** and **75.0% Stability** on CIFAR-10, outperforming baselines via **Logit Masking** and **Gradient Accumulation.**

IV. TECHNICAL SKILLS

Programming Languages: Python, C++, Java, JavaScript.
ML/DL Core: PyTorch, TensorFlow, Scikit-learn, Model Optimization & Tuning.
N**LP:** RAG, NER, Sentiment Analysis, Text Classification, TF-IDF, Word Embeddings.
Computer Vision: Object Detection (YOLO, R-CNN), Image Classification, Object Detection, Localization and Instance Segmentation, Feature Extraction.
Reinforcement Learning: DDPG, Actor–Critic, Continual Learning.
MLOps & Deployment: Docker, GitHub Action (CI/CD), FastAPI, Flask, Django, HuggingFace Spaces.
Databases & Tools: MongoDB, SQL Server.
Tools: Git, Conda, Poetry.

V. LEADERSHIP & EXTRACURRICULAR ACTIVITIES SKILLS

Member of the Executive Committee, Ho Chi Minh Communist Youth Union – Faculty of CSE.
Contributed to faculty-level student initiatives and community development programs.

- **Leader**, Volunteer campaign "Xuan Tinh nguyen" (2025) – Faculty of CSE
- **Vice Leader**, Volunteer campaign "Mua he xanh" (2024) – Ho Chi Minh City Front, Faculty of CSE
- **Vice Leader**, CSE Summer School (2024) - Faculty of CSE

VII. PROJECTS

1. Self-driving car detect objects module (Reinforcement Learning, Deep Learning) *Jul 2025 - Present*

Project goal: Built an autonomous driving system that integrates reinforcement learning with real-time object detection for safer and adaptive navigation. *Repository*

Details:

- Applied **deep deterministic policy gradient** for continuous vehicle control with object detection modules.
- Benchmarked **YOLOv11** against **Faster R-CNN (ResNet/MobileNet)**, identifying YOLOv11 as the optimal solution for real-time inference accuracy and speed.
- Implemented a robust training pipeline with **advanced augmentation (rotation, noise, color jitter)**, **AdamW optimizer, OneCycleLR scheduler**, and **early stopping**, ensuring stable convergence.
- **Optimization Pipeline:** Engineered a modular PyTorch framework using AdamW, OneCycleLR, and advanced augmentation to ensure stable convergence and reproducibility.

2. Food image classification with large label space (Computer Vision, Deep Learning) *Jul - Aug 2025*

Repository *Deployment*

Project goal: Built systematic experiments on the **Food101 benchmark (101 classes, 100K+ images)** to evaluate CNN performance under varying architectural choices, training setups, and regularization strategies.

Details:

- Compared **multiple CNN baselines**, from **shallow Conv2D** models to **SOTA** model like **EfficientNet-B2**.
- Enhanced generalization by integrating **data augmentation, early stopping, learning rate scheduling.**
- Built an evaluation pipeline for **monitoring** performance metrics, comparing models, analyzing overfitting.
- Achieved approximately **88.5% - 90% accuracy** on the benchmark.
- Developed a **Django, DRF backend** with **JWT authentication** and a **drag-and-drop** UI, containerized with **Docker** and deployed on **Hugging Face Spaces.**

3. Vietnam AQI forecasting system: Spatio-temporal regression (Deep Learning) *Nov - Dec 2025*

Repository

Project goal: Developed a unified deep learning framework using **Dual Embedding BiLSTM with Attention** to forecast multi-station Air Quality Index (AQI) across Vietnam , modeling complex spatio-temporal dependencies.

Details:

- Engineered an ETL pipeline using **Pandas**, applying **sliding windows**, **cyclical time encoding**, and **lag/rolling features** to transform raw environmental data into supervised learning inputs.
- Architected a **Dual Embedding BiLSTM with Attention** model, leveraging **Entity Embeddings** to capture spatial heterogeneity (Station/RegionID) and **bidirectional LSTM** layers to capture complex dependencies.
- Engineered a custom **weighted MSE/Huber loss** to handle outliers and implemented **gradient clipping** with **learning rate scheduling** for stable convergence.
- Developed a **recursive forecasting** strategy to extend prediction horizons to 7 days, successfully integrating the model into a web application using **Flask** and **Leaflet.js** for real-time geo-visualization and interactive analytics.
- Achieved ~**0.43 RMSE** on the test set, outperforming isolated station-based modeling approaches.

4. Agentic RAG framework for complex document retrieval *Nov - Dec 2025*

Repository *Deployment*

Project goal: Developed a scalable **RAG** system to query and extract highly accurate facts from a massive corpus of over **20,000** unstructured, OCR-degraded legal documents.

Details:

- Designed a scalable retrieval pipeline handling 256,000 vector chunks using **FAISS,** *bge-large-en-v1.5* embeddings.
- Implemented an agentic workflow with **LangChain**, **LangGraph** (featuring Grader and Query Rewriter nodes) powered by Llama-3.1 via **Groq API** to prevent hallucination and enforce factual accuracy.
- Deployed the inference engine via a **JWT-secured Django REST API**, orchestrating an **automated CI/CD** pipeline with GitHub Actions on Hugging Face Spaces.

VIII. CERTIFICATES

- **Machine Learning A-Z: AI, Python & R + ChatGPT Prize [2025]** *June 2024 - May 2025*

Platform: *Udemy* | ***Company:*** *SuperDataScienceTeam* | ***Status:*** *Completed*
Professional Certificate - Completed **100%** of 47-hour Machine Learning fundamentals implementation.

- **TensorFlow for Deep Learning Bootcamp** *Dec 2024 - June 2025*

Platform: *Udemy* | ***Company:*** *ZTM* | ***Status:*** *Completed*
Professional Certificate - Completed **100%** of 62.5-hour comprehensive training in NNs implementation.

- **The Complete Python Bootcamp From Zero to Hero in Python** *Dec 2024 - Mar 2025*

Platform: *Udemy* | ***Status:*** *Completed*
Professional Certificate - Advanced core Python programming concepts through coding exercises, projects.

- **LLM Engineering: Master AI, Large Language Models & Agents** *Aug 2025 - Present*

Platform: *Udemy* | ***Company:*** *Ligacy Team* | ***Status:*** *In progress (50%)*
Professional Certificate - build, deploy LLM applications while mastering generative AI, RAG, QLoRA, AI agents.

- **AI Engineer Agentic Track: The Complete Agent & MCP Course** *Nov 2025 - Present*

Platform: *Udemy* | ***Company:*** *Ligacy Team* | ***Status:*** *In progress (30%)*
Professional Certificate - build Agentic AI system, RAG pipeline, QLoRA fine-tuning, deploying GenAI application.

IX. LANGUAGES

Vietnamese: Native.
English: TOEIC L&R: 855 - capable of conducting technical discussions, writing clear documentation, and collaborating effectively in cross-cultural environments.