Pham Ngoc Dung

0792667152 | Education Contact | LinkedIn | Portfolio | GitHub

Education

University of Science, VNU-HCM — Ho Chi Minh City, Vietnam
Bachelor of Information Technology, CGPA: 3.95/4.0 — *Sep. 2023 – Present*

Thoai Ngoc Hau High School for the Gifted — An Giang
Physics Specialized, GPA: 9.8/10.0 — *Sep. 2020 – May 2023*

Research Interests

Generative Modeling: Diffusion models, diffusion-based approach for few-shot generation and style-content disentanglement.
Large Language Models: Looped language models, diffusion-based large language models (LLaDA).
Reinforcement Learning: Sample efficiency, exploration-exploitation trade-offs, diffusion-based approach for policy learning and trajectory optimizations in model-based reinforcement learning.

Experience

Undergraduate Research Assistant — Oct 2025 – Present
VNUHCM - University of Science, Computational Linguistics Center — *Ho Chi Minh City, Vietnam*

- Present technical research progress and proposed methodological advancements during weekly seminars with the Principal Professor.
- Collaborate with faculty and graduate researchers to brainstorm and refine architectural optimizations.
- Maintain a dedicated research commitment of 20+ hours per week, balancing high-level assistantship responsibilities with a 3.95/4.0 CGPA curriculum.

Dedicated Projects

Font Diffusion | *Python, Git, PyTorch, diffusers, Wandb* — **Code** | Oct 2025 – Present

- Engineered an end-to-end machine learning pipeline for high-fidelity dataset generation and model training, utilizing Weights and Biases for granular experiment tracking and reproducibility.
- Integrated FSTDiff innovations and developed custom modules using Skeleton Distance Transformation and Discrete Fourier Transform (DFT) to effectively decouple content features from style features.
- Designed and implemented a novel loss function (identity mapping) to bolster content-invariance, which significantly streamlined the style transformation learning process.
- Streamlined research collaboration and version control by publishing curated datasets and model checkpoints to the Hugging Face ecosystem.

Ouro Trace | *Python, Git, PyTorch, transformers, Wandb, lm-evaluation-harness* — **Code** | Oct 2025 – Dec 2025

- Benchmarked the performance and behavioral dynamics of Ouro Looped Language Models based on recent ByteDance research using reasoning tasks, **lm-evaluation-harness** and experimental tracking with **Wandb**.
- Engineered dedicated generation controls by implementing custom system prompts, strict stop sequences, and parameter tuning to ensure consistent and accurate answer extraction via regular expression parsing.
- Conducted comparative analysis between looped architectures and traditional language models , evaluating them via perplexity benchmarks and reasoning-intensive datasets.
- Discovered critical insights into the relationship between recurrent steps and reasoning performance, identifying the optimal recurrent steps value to maximize model performance on reasoning tasks while using resources efficiently.

Energy RL | *Python, Git, PyTorch, Stable Baselines 3, Gymnasium, NumPy, Numba* — **Code** | Sep 2025 – Nov 2025

- Developed a Reinforcement Learning agent to optimize network resource management under non-stationary energy demands.
- Architected a custom environment simulator using Gymnasium and NumPy, later implementing Numba to accelerate processing speeds for high-performance experience collection.
- Evaluated the efficacy of PPO and SAC algorithms by implementing custom callbacks for real-time metric logging, automated evaluation, and checkpointing.
- Achieved a performance score of 60/100, demonstrating significant improvements against random baselines even when operating under significant computational and hardware constraints.

Personal Portfolio | *Ruby, Beautiful Jekyll* **Link** | Apr 2025 – Present

- Architected and maintained a professional technical portfolio using Ruby and the Beautiful Jekyll framework to showcase personal achievements.
- Authored several technical blog posts for documenting learning journeys and sharing personal experience when working with deep learning projects.

Achievements

Flextrack Challenge 2025 (University of Wollongong) | *Rank 33rd Overall* Oct 2025
Developed a classification model (track 1) and regression model (track 2) for Detecting Energy Flexibility in Buildings; hosted on AICrowd (Username: **zoltraak**).
Leaderboard: https://www.aicrowd.com/challenges/flextrack-challenge-2025/leaderboards

Odon Vallet Scholarship (Rencontres du Vietnam) | *Recipient* 2022, 2023
Awarded for exceptional academic performance; recognized as top 0.05% among students of the same grade in the school.

Vietnam National Physics Olympiad | *Third Prize & Honor Prize* 2022, 2023
Nationwide competition for theoretical physics knowledge and understanding.

30/4 Traditional Olympiad on Physics | *Gold Medal (Rank 4th)* Apr 2021
Achieved 4th place among elite students of the same grade in Physics from Southern Vietnam.

Technical Skills

Deep Learning: PyTorch, Hugging Face ecosystem, transformers, diffusers, accelerate, Stable Baselines 3, Farama Foundation Gymnasium
Data Science & Numerical Computing: NumPy, Pandas, Numba, Einops
MLOps: Weights and Biases (Wandb)
Programming: Python, C++, Git, GitHub, Docker, VS Code, Visual Studio
Mathematics Background: Linear Algebra, Calculus I & II, Probability & Statistics
Languages: Vietnamese (Native), English (Fluent), German (Beginner)