LE VAN TRUONG

AI RESEARCHER / DATA SCIENTIST

lvtruong@selab.hcmus.edu.vn
+84 343 711 124 Ho Chi Minh City
van-truongle lee-vtruong

AI researcher and data scientist focused on turning complex, real-world data into reliable machine-learning systems. Experienced in end-to-end data pipelines, statistical evaluation, graph learning, information retrieval, and computer vision. Author of 12 research papers and an Honors IT student with a **9.50/10 GPA**.

Education

Vietnam National University, Ho Chi Minh City — University of Science Sep 2023 – Present
Honors Program in Information Technology; expected graduation: 2027 **GPA 9.50/10**

Selected coursework: Probability and Statistics (10.0), Calculus 2B (10.0), Discrete Mathematics (9.7), Applied Mathematics and Statistics (9.6), Linear Algebra (9.3).

Hung Vuong High School for the Gifted, Gia Lai 2020 – 2023
Specialized in Computer Science **Grade 12 GPA 9.7/10**

Research & Professional Experience

Research Member | VNG - Saigon AI Hub, Applied AI Lab Batch 1 (VR Story Train AI) Nov 2025 – Oct 2026

- Member of one of 10 selected teams in Saigon AI Hub's inaugural Applied AI Lab; developed **CrisisX**, a mobile VR clinical-communication training system.
- Integrated Speech-to-Text, Gemini 3.1 Flash-Lite, multimodal retrieval, VAD emotion modeling, and adaptive agents to generate context-aware dialogue, scenario escalation/de-escalation, and post-session communication and empathy assessment.
- Prototype evaluation with 14 healthcare-related students achieved **UMUX-LITE 76.8/100**; designed for smartphone + Google Cardboard deployment. Project outcomes include *Lumina-HT2* (ACM ICMR 2026) and *CrisisX* (IEEE ISMAR 2026 poster).

AI Intern | Global Tech Jul 2025 – Jan 2026

- Built a Vietnamese information-retrieval and RAG pipeline over **100K news articles** using **hiieu/halong_embedding**, **Milvus (IVF-SQ8)**, and **PhoRanker**, improving **Hits@1 from 41% to 47%** (**+6 pp; +14.6% relative**) over raw dense retrieval on a 100-query gold-standard benchmark.
- Developed NLP components for **named-entity recognition (NER), retrieval, and text-based question answering**; optimized retrieval and reranking to achieve **94.5% Hits@20** at approximately **0.174 s/query** on a 200-query evaluation set.

AI Project Contributor | DIGIIN (Digital Insurance Services JSC) Mar 2026 – Aug 2026

- Contributed to an AI-powered digital loss-assessment solution using computer vision to assess damage and detect potential fraud, accelerating the workflow from field inspection to automated claim settlement.

Selected Data Science & AI Projects

ChronoSeek | Temporal-First Agentic Lifelog Retrieval ML / Data Science Researcher, 2026
Python, Multimodal Retrieval, RAG, Neo4j, Knowledge Graphs, LLMs

- **Problem:** retrieve evidence for complex temporal questions from a large, unstructured lifelog collection.
- **Approach:** transformed **725,000+ images** into approximately 4,000 temporal events; built an Event Knowledge Graph and a temporal-first pipeline combining semantic retrieval, event relations, reasoning, and RAG.
- **Results / baseline evaluation:** benchmarked against flat semantic-retrieval baselines; reached **Keyword Hit@10 ≈ 0.99**, complete QA coverage, and approximately **6.7 s p95 latency**.
- **Application:** reusable architecture for enterprise knowledge search, customer-support Q&A, and time-aware document retrieval; accepted at LSC 2026, ACM ICMR 2026.

ResiliRoad | Spectral-Prior Graph Learning for Road-Network Resilience Data Scientist / ML Researcher, 2026
Python, PyTorch, GNNs, OpenStreetMap, Network Analysis, Bootstrap Evaluation

- **Problem:** estimate road-network connectivity loss quickly when disruptions make exact recomputation costly.
- **Approach:** built graph-data pipelines over **13 OpenStreetMap areas across 6 countries**; combined analytical spectral priors with residual GNNs and tested GCN, GraphSAGE, and edge-aware MPNN models.
- **Results / baseline:** reduced MAE from **0.096 to 0.057** versus analytical and non-graph baselines; reduced scenario evaluation from approximately **1.30 s to 23 ms** on graphs up to **20,000 nodes**.
- **Application:** supports resilient logistics, supply-chain planning, disruption response, and route-prioritization decisions.

SKIA | Motion-Robust Remote Photoplethysmography Machine Learning Researcher, 2026
Python, PyTorch, OpenCV, NumPy, SciPy, Signal Processing, Statistical Testing

- **Problem:** estimate physiological signals reliably from facial video when lower-face motion corrupts remote photoplethysmography.
- **Approach:** designed a multi-stream model combining raw appearance, lower-face suppression, and temporal differences; built preprocessing, ablation, runtime, frequency-domain, and statistical-testing pipelines on UBFC-rPPG.
- **Results / baseline:** improved CHROM correlation from **0.3507 to 0.4166** and HR MAE from **33.50 to 23.44 BPM**; full SKIA achieved **0.7522 correlation** and **3.73 BPM HR MAE**.
- **Application:** enables more robust contactless vital-sign monitoring for telehealth and smart-care workflows; accepted at ICMV 2026.

Technical Skills

Programming & Data	Python SQL C/C++ Pandas NumPy SciPy Git Linux
ML & Modelling	PyTorch, scikit-learn, graph neural networks, deep learning, NLP, computer vision, embeddings, RAG, knowledge graphs, time-series and signal processing.
Analytics & Engineering	Experimental design, hypothesis testing, bootstrap evaluation, feature engineering, model evaluation, ablation studies, error analysis, data visualization, Docker, Neo4j.

Selected Publications

- ***"Enhancing Multilingual Fact-Checking Retrieval with Ensemble Bi-Encoders and Architecture-Aware Hard Negative Reranking."*** Le, V. T., Nguyen-Tran, D. M., Huy, N. T., & Le, T. (2026). **KES 2026 - Scopus B Conf.**
- ***"ChronoSeek: Temporal-First Agentic RAG with Event Knowledge Graph for Lifelog Retrieval."*** Le, V. T., Tran, M. T., & Huynh, V. T. (2026). **ACM ICMR 2026 Workshop LSC - Scopus B Conf.**
- ***"SKIA: Spatial-Knowledge Informed Attention for Robust Remote Photoplethysmography under Lower-Face Motion Artifacts."*** Le, V. T., Tran, M. T., & Huynh, V. T. (2026). **ICMV 2026 - Scopus C Conf.**
- ***"From Benchmarking to Reasoning: A Dual-Aspect, Large-Scale Evaluation of LLMs on Vietnamese Legal Text."*** Le, V. T., et al. (2026). **FJCAI 2026 - National AI Conf.**
- ***"A Two-Stage, Object-Centric Deep Learning Framework for Robust Exam Cheating Detection."*** Le, V. T., et al. (2026). **FJCAI 2026 - National AI Conf.**
- ***"FLeX: Fingerprint Leakage Evaluation across Super-Resolution and Cross-Domain Matching."*** Le, V. T., et al. (2026). **FDSE 2026 - Springer CCIS, Scopus.**

Awards, Certifications & Leadership

Scholarships	**Odon Vallet Scholarship** (2026); Pony Chung Foundation and VPBank Future Banker Scholarships (2025–2026); 100% Tuition and Merit Scholarships, HCMUS (2023–2026); VietSeeds Foundation Scholar (2023–2025); KoiBi, FPT, VEF, SAC, and alumni-funded awards.
Research & Academic	Seven conference papers; presenter, 3rd HCMUS Young Scientists Conference (2026); HCMUS ICPC Team (2024); Faculty of IT Golden Achievement Board (2024–2025); "Read to Share" Champion (2025) and Consolation Prize (2026).
Innovation & Awards	Consolation Prize, Ho Chi Minh City Youth Creative Ideas Competition (2025); First Prize, Gia Lai Provincial Excellent Student Contest in Informatics (2022); Third Prize, Hue-ICT Challenge Pro and Regional National Youth Informatics Contest (2023).
Training	Machine Learning Specialization (DeepLearning.AI); SQL Advanced (HackerRank); Responsible AI, LLM, and Vertex AI training (Google); Computer Vision, ML Explainability, and AI Ethics (Kaggle).
Leadership & Service	Champion, HCMUS Student Leadership Competition (2023); Champion, Faculty of IT Leadership Competition (2024); organizer of AI talks and academic-support programmes; Green Summer volunteer and ten-time blood donor (2024–2026).
Languages	Vietnamese (native); English: VSTEP Level 4 / CEFR B2, 7.0/10; technical writing, research presentations, and international collaboration.

Global Experience & Community Service

- **Sole Vietnam National University system delegate**, ASEAN University Network Summer Camp, De La Salle University, Philippines (2026); participant, NUS–AUN Asian Undergraduate Symposium and REGEN Asia Summit, National University of Singapore (2025).
- **Community service:** Green Summer Campaign; Tan Binh Shelter outreach; incoming-student orientation support; and ten voluntary blood donations from 2024 to 2026.