

LE THANH THIEN

(+84) 961 941 329 | thienlt.7@gmail.com | thanhthienle.github.io

EDUCATION

Hanoi University of Science and Technology, Hanoi	Aug 2019 – Aug 2023
<i>Bachelor of Science in Computer Science, Distinction equivalent</i>	
• Achievements	Talent Scholarship 2019 (merit-based, value of 1,300 USD)
• Thesis	Parameter-efficient Continual Relation Extraction , combining generative replay with a novel multi-objective optimization method for competitive Continual Relation Extraction (GPA 4.0)
• Projects	Volatility Prediction for Vietnam Stock Market ↗ , benchmarking ARCH, GARCH, SVM, and Multi-Layer Perceptron on VNINDEX and VN30INDEX Data Platform for Financial Market ↗ , building big data pipeline for the financial market Bird Classification , training ResNet models, resulting in accuracy of over 90% Image Captioning ↗ , building and training Show and Tell model using CNNs and LSTMs
• Modules	Artificial Intelligence (GPA 4.0), Computer Architecture (GPA 4.0), Data Science (GPA 4.0), Evolutionary Computing (GPA 4.0), Information Retrieval (GPA 4.0), Linear Algebra (GPA 4.0), Machine Learning and Data Mining (GPA 4.0), Programming Techniques (GPA 4.0)

VNU.HCM High School for the Gifted, Ho Chi Minh city	Aug 2016 – May 2019
<i>High School Diploma (GPA 9.7)</i>	
• Achievements	Odon Vallet Scholarship 2019, value of 430 USD, from Rencontres Du Vietnam ↗

EXPERIENCE

VinAI Research, Hanoi AI Research Resident	Feb 2023 – Present
<i>Top 20 global company for leading AI research in 2022</i>	
• Research cross-lingual information extraction • Conduct studies on continual learning in information extraction, resulting in 3 papers • Research problems related to Large Language Models, resulting in 1 paper • Collect and preprocess data to train chat Large Language Model for Vietnamese language • Trained and benchmarked Automatic Speech Recognition models on public/private data for Vietnamese language	

VND Credit, Hanoi Data Science Engineer	Jul 2022 – Feb 2023
<i>P2P lending platform</i>	
• Wrote, trained, and deployed API for job-matching model in JOBO job recommendation system • Gathered and standardized data on job opportunities for training AI models extracting useful information from job description, achieving precision and recall of over 90% • Collaborated with 6 engineers and 3 interns to build graph neural networks for credit scoring from mobile-usage data	

HUST Data Science Lab (BKAI), Hanoi Research Student	Apr 2020 – Aug 2023
<i>Affiliated with NAVER – most-used search portal in South Korea</i>	
• Researched deep generative models (e.g. GANs, Diffusion Models) and theoretical generalization for deep neural networks, led by Assoc. Prof. Than Quang Khoat • Conducted project “Deep learning for imbalanced data”, designing architectures and training paradigms for deep neural networks to enhance performances on skewed and long-tailed data	

ACHIEVEMENTS

Physics Unlimited Explorer Competition (PUec) Honorable Mention	2018
<i>International Physics competition originated from Princeton University</i>	
• Track: Statistical Mechanics	
Vietnam Physics Olympiad (VPhO) Second Prize	2019
<i>Most prestigious Physics competition for high school students in Vietnam.</i>	
• Top < 0.05% of nationwide high-school students in 2018-2019 school year	

PUBLICATIONS

Asterisk (*) denotes equal contribution

Continual Relation Extraction via Sequential Multi-Task Learning
Thanh-Thien Le*, Manh Nguyen*, Tung Thanh Nguyen*, Linh Van Ngo and Thien Huu Nguyen, AAAI 2024, Vancouver, Canada

SharpSeq: Empowering Continual Event Detection through Sharpness-Aware Sequential-task
Thanh-Thien Le*, Viet Dao*, Linh Van Nguyen*, Thi-Nhung Nguyen, Linh Van Ngo and Thien Huu Nguyen, NAACL 2024, St. Julian's, Malta

Realistic Evaluation of Toxicity in Large Language Models
Tinh Son Luong*, Thanh-Thien Le*, Linh Van Ngo and Thien Huu Nguyen, ACL 2024, St. Julian's, Malta

PhoWhisper: Automatic Speech Recognition for Vietnamese
Thanh-Thien Le, Linh The Nguyen, Dat Quoc Nguyen, Tiny Paper @ ICLR 2024, Vienna, Austria

ToVo: Toxicity Taxonomy via Voting
Tinh Son Luong*, Thanh-Thien Le*, Thang Viet Doan*, Linh Ngo Van, Thien Huu Nguyen and Diep Thi-Ngoc Nguyen, under review

Lifelong Event Detection via Optimal Transport
Viet Dao*, Van-Cuong Pham*, Quyen Tran*, Thanh-Thien Le*, Linh Ngo Van, Thien Huu, under review

Few-Shot, No Problem: Descriptive Continual Relation Extraction
Thanh Nguyen*, Anh Duc Le*, Quyen Tran*, Thanh-Thien Le*, Linh Ngo Van, Thien Huu, under review

Adaptive Prompting for Continual Relation Extraction: A Within-Task Variance Perspective
Minh Le, Tien Ngoc Luu, An Nguyen The, Thanh-Thien Le, Thien Trang Nguyen Vu, Tung Thanh Nguyen, Linh Van Ngo and Thien Huu Nguyen, under review

SKILLS	
• Programming	Python (NumPy, pandas, Matplotlib, PyTorch, PySpark, Hugging Face Transformers), C++
• Machine Learning	Linear Regression, K-Means, SVMs, KNN, Probabilistic Models, DNNs, GNNs, LLMs
• Mathematics	Calculus, Linear Algebra, Probability and Statistics, Optimization, Discrete Mathematics
• Languages	English (Fluent), Vietnamese (Native)

REFERENCES	
Asst. Prof. Thien Huu Nguyen	thien@cs.oregon.edu https://ix.cs.uoregon.edu/~thien/
Linh Van Ngo, Ph.D.	linhnv@soict.hust.edu.vn https://users.soict.hust.edu.vn/linhnv
Assoc. Prof. Khoat Than	khoattq@soict.hust.edu.vn https://users.soict.hust.edu.vn/khoattq