

THANH LONG LY

 lytlong.pers@gmail.com

 @tlong-ds

 Thanh Long Ly

 +84 868 421 575

Education

2023 – 2027

 **National Economics University (NEU)**, Hanoi, Vietnam.

Major: Data Science in Economics and Business

CGPA: 3.55/4.0

Relevant coursework: Programming, Database Management System, Discrete Math, Econometrics, Data Structures & Algorithms, Data Preparation and Visualization, Machine Learning.

Research Experience

2025 – Present

 **Research Assistant**, Predictive Models and Machine Learning Research Applications, NEU.

- Designed validation frameworks for predictive economic models, conducting **ablation studies** to isolate feature importance and ensure statistical robustness.
- Fine-tuned transformer-based architectures (BERT/RoBERTa) for domain-specific NLP tasks, monitoring **loss convergence** and optimizing inference latency for real-time deployment.
- Key Contribution:** Engineered a robust data preprocessing pipeline that reduced text noise by 20%, significantly aligning with downstream model requirements.

2024 – 2026

 **Student Researcher**, Scientific Student Research Competition, NEU.

- Co-authored the paper "*The Regulatory Role of Provincial Governments in the Impact of Labor Shift on Agricultural Production Efficiency in Vietnam*", published in the *Proceedings of the National Scientific Conference 2024*.
- Implemented multivariate regression models to quantify policy impact, utilizing rigorous statistical tests (e.g., F-test, t-test) to validate hypothesis significance.
- Recognized with the university-level "**Student Research Award 2026**" for research methodological excellence.

Research Publications

Projects

2025 – Present

 **RAG-based Information Retrieval for E-commerce** (Rang Dong).

- Engineered a Retrieval-Augmented Generation (RAG) pipeline utilizing **Qdrant** for dense vector retrieval and **LLM** re-ranking.
- Optimized semantic search embeddings for Vietnamese, improving **Recall@10** by 18% compared to baseline BM25 keyword search.

Projects (continued)

December, 2025

Domain-Specific Conversational AI (Gigabyte Hackathon).

- Designed a low-latency dialogue system for hardware support, implementing strict prompt grounding to minimize hallucination.
- Benchmarked context retention strategies, achieving a stable **response latency under 800ms** for complex multi-turn queries.
- **Result:** 1st Place for architectural robustness and inference efficiency.

November, 2025

Time-Series Forecasting for Meteorological Data.

- Developed an ensemble regression framework (CatBoost + XGBoost) to forecast hyper-local weather patterns.
- Conducted ablation studies on feature sets, reducing **RMSE by 15%** and achieving an **F1-score of 0.89** on extreme weather event classification.

Working Experience

2025 – Present

AI Engineer Intern, Alpha Data Academy, Hanoi, Vietnam.

- Developed an Agentic RAG System for Rang Dong, a company in the lighting industry.
- Participated in the design of data ingestion pipeline for the system.
- Optimized the response to be natural and human-like.

Honors, Awards & Outreach

Oct. 2025 – Dec. 2025

Champion (1st Place), Master AI, Master Career Hackathon.

- Developed a high-precision QA system for technical hardware specifications, overcoming challenges in terminology ambiguity and context retention.
- Evaluated by industry judges as the most technically robust solution among 20+ participating teams.

Skills

Research Stack

Languages: Python, C++, SQL.

ML/AI: PyTorch, TensorFlow, HuggingFace Transformers, LangChain, Scikit-learn.

Engineering

Backend/DB: Qdrant (Vector DB), PostgreSQL, MongoDB, Supabase.

DevOps/Tools: Docker, Kubernetes, AWS, Git, MLFlow.

Languages

English (IELTS 7.0), Vietnamese, Chinese.