

# 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

### Conference Proceedings

1

M. D. Phung, B. D. T. Vu, T. L. Ly, T. H. T. Duong, and K. L. Pham, "The regulatory role of provincial governments in the impact of labor shift on agricultural production efficiency in Vietnam," in *Proceedings of the National Scientific Conference: Vietnam's Economy in 2024 and Prospects for 2025, Promoting Economic Institutional Reform in the New Context*, Hanoi, Vietnam: National Economics University Publishing House, 2025, pp. 488–501.

### In Progress

1

T. P. Trinh, A. T. Tran, T. L. Ly, A. K. Phan, Q. A. Bui, and N. H. Dang, *Agentic B2B2C AI - how AI has been transforming B2B2C e-commerce in vietnam*, In Preparation, 2026.

## Projects

2025 – Present

### [R] RAG-based Information Retrieval for E-commerce (Rang Dong).

- Designed a **hybrid retrieval system** (dense + sparse + RRF reranking) on **Qdrant**.
- Built data ingestion pipelines and LLM response postprocessing to match business intents; deployed on **Render** with **Opik** monitoring and **MongoDB** logging.

December, 2025

### [R] Domain-Specific Conversational AI (Gigabyte Hackathon).

- Architected a **hybrid agentic system** (LangGraph + Retion.ai) with specialized REST tools (ROI calculation, review search) to handle complex B2B/B2C scenarios.
- **Result:** Awarded **1st Place** for the comprehensive scenario design and robust system performance.

November, 2025

### [R] Time-Series Forecasting for Meteorological Data.

- Developed a **29-model framework** (CatBoost/LightGBM/XGBoost) with 60+ engineered features, optimized via Optuna (2,400+ trials).
- Achieved **RMSE 1.52°C / R² 0.902** (daily) and **RMSE 1.61°C / R² 0.907** (hourly), deploying the pipeline via FastAPI and ClearML.

## Working Experience

2025 – Present

### [R] 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

### [R] 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

### [R] Languages: Python, C++, SQL.

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

Engineering

### [R] Backend/DB: Qdrant (Vector DB), PostgreSQL, MongoDB, Supabase.

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

Languages

### [R] English, Vietnamese, Chinese.