

THU PHUONG NGUYEN

Email: phuongnt@unist.ac.kr

Webpage: <https://phuongnt197.github.io/>

EDUCATION

Ulsan National Institute of Science and Technology (UNIST)

Department of Computer Science and Engineering
Cumulative GPA: 4.08/4.3

Ulsan, Korea

Sep 2024 - present

Ulsan National Institute of Science and Technology (UNIST)

Department of Computer Science and Engineering
Cumulative GPA: 3.83/4.3 (Magna Cum Laude)
UNIST Scholarship and Global Dream Scholarship that covers full tuition fee and meals for 4 years. Summer 2022 Exchange Student at PolyU (Hong Kong).

Ulsan, Korea

Sep 2020 – Aug 2024

HONORS AND AWARDS

Outstanding Student Awards of the Semesters in 2020, 2023, 2024	2024
Bronze Medal in the 2021 U-Challenge Festival	2021
Rector's award of merit for Outstanding Academic Excellence	2019
Outstanding Young Face Awards in School Year of 2016 - 2018	2018
Third Prize in the Vietnam Science and Engineering Fair 2017 - 2018	2018
Third Prize in the 2017 Robot Creativity Contest	2017

RESEARCH EXPERIENCE

Interactive Multimodal Machine Learning Lab, UNIST

Graduate Research Assistant
Supervisor: Professor Taehwan Kim
Project: VEHME - A Vision-Language Model for Evaluating Handwritten Mathematics Expressions

Ulsan, Korea

Sep 2024 – Present

- Designed dual-phase pipeline (SFT+RL) with a novel expression-aware visual prompting module that improved performance of open-source VLMs by 30%.
- Achieved SOTA performance on the AIHub and FERMAT benchmarks—leading among open-source models and coming close to GPT-4o and Gemini-2.5.

Software Testing and Analysis Research Lab, UNIST

Research Intern
Supervisor: Professor Mijung Kim
Project: SWE-Agent: Supporting Software Engineering Tasks for Java using Large Language Models with Agents.

Ulsan, Korea

Jan 2023 – Feb 2024

- Addressing challenges in automated high-quality unit tests generation for bug reproduction, finding and fixing buggy methods for fault localization and program repair tasks based on bug reports.
- Created an environment where we define agents using Langchain, incorporating an efficient prompt for Large Language Models integration, which outperforms previous method with a remarkable 63.8% improvement in acc@1 for bug reproduction and by 1.6 times in pass@1 for program repair in the Defects4j benchmark.

Project: Generating Unit Tests VTW-DnA Low Code Platform

- Implemented automated testing using AFLGo mutation-based fuzzing techniques, achieving a notable time improvement of 2 times compared to coverage-based greybox fuzzing (CGF).
- Developed a versatile testing framework facilitating multi-language unit test generation for Python and JavaScript program generated by DnA platform, powered by a Large Language Model.
- Financially supported by VTW company.

Sustainable Structural Systems and Materials Lab, UNIST

Research Intern

Ulsan, Korea

Jun 2021 – Jul 2022

Supervisor: Professor Myoungsu Shin

Project: Develop machine learning models to predict the seismic vulnerability of building structures.

- Developed machine learning models that can predict the seismic responses of planar steel moment-resisting frames subjected to ground motions using artificial neural network (ANN) and extreme gradient boosting (XGBoost). The method achieves high R^2 score of up to 96.1%.
- Created a graphical user interface (GUI) tool using PysimpleGUI for the preliminary estimation of the seismic drift responses of steel moment frames based on the trained XGBoost model.
- Financially supported by 2021 Brain to Society Industry-Friendly Research Project.

Radio Physics Lab, Hanoi University of Science

Research Intern

Hanoi, Vietnam

May 2017 – Jun 2018

Supervisor: Professor Quang Thao Le

- Researched on the Medication Reminder for Absent-minded people Project.
- Built robots that throw the ball into the basket.
- Designed printed circuit boards and researched semiconductor components.

PUBLICATIONS

VEHME: A Vision-Language Model For Evaluating Handwritten Mathematics Expressions

Thu Phuong Nguyen*, Duc M. Nguyen*, Hyotaek Jeon, Hyunwook Lee, Hyunmin Song, Sungahn Ko** and Taehwan Kim**. In *Proceedings of EMNLP 2025: The Conference of Empirical Methods in Natural Language Processing (to appear)*.

PROJECTS

Drag-guided 3D Motion Generation | Pytorch

- Applied ControlNet training with control signals from dynamically extracted joints.
- Leveraged a VLM to generate textual description from user drags drawn on the skeleton to control motion with simple drags, achieved SOTA performance by outperforming previous approaches by 20%.

Multi-agent Multi-Destination Packet Routing Using Deep Reinforcement Learning | Python

- Developed a fully distributed multi-agent DRL framework for multi-destination packet routing
- Significant improvements in E2E delay and congestion avoidance, particularly under high traffic loads.

Algorithmic Trading Bot | Python

- Designed and implemented a novel trading strategy using RSI and Bollinger Band technical indicators.
- Deployed the system into Interactive Brokers, yielding a 0.6% return on investment daily on average.

Stock Market Analysis | Machine Learning, Time Series Similarities,

- Conducted both qualitative and quantitative analysis of S&P500 stocks data over 1 year.
- Classified the stocks into categories that are sensitive to their respective markets.
- Compared the efficiency of DTW and Euclidian metrics in K-Means clustering.
- Regressed each cluster to observe the relationship between excess and factor returns.

Handling multiplex inputs for HTTP server | C/C++, socket, cache, multiplexing

- Implementation of multiplexing handling strategy from scratch using C/C++.
- Optimized requests handling parallel socket clients.
- The system is able to serve more than 2 million requests in a span of 1 hour.

Currency Rates Exchange Prediction | PyTorch, ARIMA, LSTM

- Rendered visualizations for qualitative analysis providing insights to problem solving.
- Derived general ARIMA formula for analysis achieving a MSE score of 5.6×10^{-5} .
- Implemented LSTM-based model for forecasting achieving a MSE score of 3.5×10^{-5} .

Taxi Demand Prediction | PyTorch, Transformers, CNN

- Conducted EDA on New York Yellow Taxi Data with millions of data points.
- Optimized data processing time by 2 orders of magnitude.
- Implemented a spatial-temporal neural network based on Transformers and CNN.

Artists Identification | Python, Tensorflow, Residual Network

- Built a CNN model to identify 11 artists through their artworks.
- Achieved 93% accuracy on the test set.

EXTRACURRICULAR ACTIVITIES

The Conference on Empirical Methods in Natural Language Processing	Suzhou, China
Volunteer	Nov 4 th – 9 th , 2025
	<ul style="list-style-type: none"> • Assisted throughout the conference in multiple roles - managing plenary session logistics, supporting poster sessions, and facilitating registration operations.
The 10th Annual Conference of Vietnamese Young Scientists (ACVYS 24)	Seoul, Korea
Staff	Sep 8 th , 2024
	<ul style="list-style-type: none"> • Assisted with managing the poster presentation and tea break sessions, supporting smooth coordination among 41 presenters and hundreds of young scholars.
Vietnamese Student Association at UNIST	Ulsan, Korea
Chairwoman	Sep 2024 – Present
	<ul style="list-style-type: none"> • Organized 10 cultural and community events to promote cross-cultural engagement and support the well-being of Vietnamese students.
Teaching Assistant	Ulsan, Korea
TA	Feb 2022 – Present
	<ul style="list-style-type: none"> • Course: Introduction to AI Programming I, Theory of Computation, Discrete Mathematics, Statistical Computing. • Responsible for answering questions, and grading assignments and exams for over 150 students each semester.
2022 Vietnamese Students in Korea SW Education Program	Seoul, Korea
Scholar	Aug 2022 – Sep 2022
	<ul style="list-style-type: none"> • Hosted by the National IT Industry Promotion Agency (NIIPA) • Studied Java development and Elastic Search.

- Produced visualization using Kibana.

UNIST International Students Organization (UISO)

Ulsan, Korea

Event Manager

Feb 2022 – Dec 2022

- Organized over 15 large scale events and activities for over 200 international students.

Campus Advisor Mentorship Program

Ulsan, Korea

Mentor

Jun 2021 – Dec 2021

- Supported new international students in Korean culture exploration, UNIST Campus Life, choice of major and study abroad life.

Maths Open Day 2018

Hanoi, Vietnam

Volunteer

Nov 2018

- Organized math and science experiences "In the land of magic math" from partner units for students, mathematicians, and educators.

GreenHSGS Robotics Team

Hanoi, Vietnam

Head of Research and Planning Department

Sep 2017 – Jun 2018

- Organized a robotics training day for students at High School for Gifted Students.
- Opened booths in Exhibition Fair of Technology “CyberPiece” and Exhibition of Educational Activities of Group 91-94 Hanoi

OTHER SKILLS

Languages: English: professional working proficiency, Vietnamese: native, Chinese: limited working proficiency, Korean: professional working proficiency.

Programming languages: C/C++, Python, Scala, Java, JavaScript, LaTeX

Frameworks & libraries: TensorFlow, Pytorch, Scikit-learn, Pandas, Hugging Face

Others: Altium Designer, SolidWorks, Unix/Linux, Git/Github actions, Figma.