Minh Nam TRAN

☐ +84 382 563 952 | @ minhnam.tran820@gmail.com

in LinkedIn | ⌂ GitHub | 🌐 Portfolio | 🎓 Google Scholar

RESEARCH INTERESTS

My research interests are in **Natural Language Processing**, including (but not limited to):

- **Language Model-related Problems**: Natural language understanding, logical reasoning, hallucination, knowledge integration, trustworthiness, and robustness.
- **Applications of Language Models**: Using language models for domain-specific problems, real-world problems, or various NLP downstream tasks, such as information retrieval, question-answering, summarization, and dialogue systems.
- **Training/Learning Methods**: Learning strategies for improving the performance of language models and making them think logically like humans.

EDUCATION

University of Science, Vietnam National University - HCMC *Sep. 2020 − Oct. 2024*
Bachelor of Science in Computer Science (Advanced Program) *Ho Chi Minh, Vietnam*

- GPA: 3.98/4.00 - Degree class: Excellent - Rank: 1/90.
- Thesis title: "*Exploring and Improving Language Understanding Abilities of Vietnamese Language Models,*" advised by Assoc. Prof. DINH Dien and Dr. NGUYEN Hong Buu Long (score: 4.0/4.0).

SELECTED PUBLICATIONS

1. **Minh-Nam Tran**, Phu-Vinh Nguyen, Long Nguyen, and Dinh Dien. 2024. ViGLUE: a Vietnamese general language understanding benchmark and analysis of Vietnamese language models. In *Findings of the Association for Computational Linguistics: NAACL 2024*, 4174–4189. % DOI: 10.18653/v1/2024.findings-naacl.261.

2. Tuan-An To, **Minh-Nam Tran**, Trong-Bao Ho, Thien-Loc Ha, Quang-Tan Nguyen, Hoang-Chau Luong, Thanh-Duy Cao, and Minh-Triet Tran. 2024. Multi-perspective traffic video description model with fine-grained refinement approach. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 7075–7084. % DOI: 10.1109 /CVPRW63382.2024.00701.

EXPERIENCE

Computational Linguistics Center, University of Science *May 2022 − Present*
Undergraduate Research Assistant *Ho Chi Minh, Vietnam*

- Research under supervision of Prof. DINH Dien and Dr. NGUYEN Hong Buu Long.
- Organize in-lab seminars about Transformer and language models (BERT, GPT, etc.).
- Complete my project to evaluate Vietnamese language models and publish a Vietnamese general language understanding evaluation benchmark (ViGLUE) with findings about the pre-trained language models under few-shot learning and direct fine-tuning (code).

- Design a dual relationship learning method for improving language models on the natural language inference task in the Vietnamese medical domain and outperformed the baseline methods on the ViMedNLI and ViNLI$_{health}$ datasets.

Viettel Group *Apr. 2023 − Oct. 2023*
AI Intern, Viettel Digital Talent 2023 *Hanoi, Vietnam*
- Completed two projects in machine learning/deep learning under the supervision of Dr. Nguyen Van Nam: the table-to-text challenge and the text-to-SQL problem.
- Fine-tuned encoder-decoder Transformer models, including BART and T5, for table-to-text generation with synthesized data created by gpt3.5.
- Applied the QLoRA method to fine-tune CodeLLama-7B and CodeLlama-14B models on a single H100 GPU for the text-to-SQL problem, achieving the performance at 72.7% execution accuracy and 61.5% exact match on the test subset of Spider benchmark.
- Deployed text2sql models under 4-bit quantization with FastAPI, Docker, and `Llama-cpp`.

Users and Information Lab, KAIST *Jun. 2023 − Aug. 2023*
Visiting Research Student *Daejeon, South Korea*
- Worked under the supervision of Prof. Alice Oh and PhD. student Rifki Afina Putri.
- Investigated adapting the BLOOM model family for four Vietnamese downstream tasks using low-rank adaptation and multitask instruction tuning (code).
- Studied the length-penalized loss to help language models focus on short output tasks.

Honors & Awards

Certificate of Merit for Outstanding Students in Research Activities *Aug. 2024*
- First author of the publication titled "ViGLUE: a Vietnamese General Language Understanding Benchmark and Analysis of Vietnamese Language Models."

Second Prize of the 4th AI Challenge Ho Chi Minh City 2023 *Oct. 2023*
Theme: "Event Retrieval from Visual Data"
- Competed with 340 teams in the qualifying round and 44 teams in the final round.
- Designed an efficient multi-modal retrieval system searching images over a million photos based on a textual query or a sample image with the similarity search method using the Faiss vector database.
- Applied CLIP and BLIP image-text models to extract image embedding vectors from keyframes at specific timestamps of videos and utilized Whisper to extract transcripts from videos.

Full-ride Academic Encouragement Scholarship 2021-2022 Academic Year *May 2023*
- Highest GPA among a cohort of approximately 90 students in the Advanced Program in Computer Science during the 2021-2022 academic year.

Half-fee Academic Encouragement Scholarship 2020-2021 Academic Year *Jun. 2022*
- Top three students with the highest GPA among approximately 90 students in the Advanced Program in Computer Science during the 2020-2021 academic year.

Presentations

NAACL 2024 Conference, Mexico City, Mexico *Jun. 2024*
- Online poster presentation: "ViGLUE: A Vietnamese General Language Understanding Benchmark and Analysis of Vietnamese Language Models."

CVPR 2024 Workshop, Seattle, US *Jun. 2024*
- In-person oral presentation: "Multi-perspective Traffic Video Description Model with Fine-grained Refinement Approach."

Training & Relevant Coursework

TensorFlow Developer Certificate *Jul. 2023*
by TensorFlow

Deep Learning *Jun. 2022*
by DeepLearning.AI

Community & Volunteer Work

AI Viet Nam *Jul. 2022 – Aug. 2023*
Volunteer *Online Community*
- Led a team of six members in a code summarization project for developing a simple but highly interactive Streamlit application to generate docs from Python code ([demo link](#)).
- Worked in the Daily Math Team to create math exercises for ML/DL and translated TensorFlow deep learning projects into PyTorch code to support the learning community.

Google Developer Student Club, University of Science *Sep. 2021 – Aug. 2022*
Core Team Member *Ho Chi Minh, Vietnam*
- Designed content for Day 2 of the "CS101" workshop series in collaboration with SAB Academy with the topic of VIM text editor and providing technical resources.
- Co-organized "First Step To PM Era" with AIESEC, holding the knowledge-sharing session with industry professionals to guide students towards product management careers.

Skills & Technologies

- **Programming:** Python, C/C++, Java, R, Shell, JavaScript.
- **Frameworks:** NumPy, PyTorch, TensorFlow, Scikit-learn, FastAPI, llama-cpp, LangChain.
- **Tools and Technology:** Git, Docker, Linux server, Slurm, LaTex.
- **Specialized skills:** Deep Learning, Natural Language Processing, Large Language Model.

References

Assoc. Prof. DINH Dien
Director of Computational Linguistics Center
Faculty of Information Technology
University of Science, VNU-HCM
Tel: (+84) 908 278 207
Email: ddien@hcmus.edu.vn

Dr. NGUYEN Hong Buu Long
Lecturer
Faculty of Information Technology
University of Science, VNU-HCM
Tel: (+84) 985 187 289
Email: nhblong@fit.hcmus.edu.vn