

VIET H. PHAM

☎ +84 964 · 217 · 704 ◇ ✉ viethungpham0304@gmail.com ◇ 🐙 GitHub

Biography: I completed my Bachelor of Science Degree in Computer Science at the [University of Science](#) in Vietnam, advised by [Assoc. Prof. Dien Dinh](#). My bachelor's thesis focused on Natural Language Processing (NLP) and Theoretical Linguistics, in which I integrated the [Abstract Meaning Representation \(AMR\)](#) semantic graph into Neural Machine Translation. After graduating, I worked in the IT industry for more than one year as an AI Engineer. Meanwhile, I am also a research assistant at the [Computational Linguistic Center \(HCMUS\)](#), advised by Assoc. Prof. Dien Dinh. My research interests include machine translation, language modeling, and QA.

EDUCATION

University of Science 2016 - 2020
B.Sc in Computer Science *Ho Chi Minh City, Vietnam*
· Thesis: Integrating Abstract Meaning Representation into Neural Machine Translation
· Defend: 10/10 (Excellent undergraduate thesis)

WORK EXPERIENCE

Auburn University August 2022 - present
Research Assistant *Auburn, AL, USA*

- Research topics: *Interpretable Machine Learning* and *Phrase Language Modelling*
- Multiple-round reading comprehension
- Build a phrase language model that can distinguish different meanings of a phrase

Computational Linguistic Center December 2019 - August 2022
Research Assistant *Ho Chi Minh City, Vietnam*

- Research interests: *Neural Machine Translation (NMT)* and *Semantic Representation*
- Build a Graph Encoder for encoding semantic information from semantic graphs
- Integrate semantic information into Neural Machine Translation system
- Co-Advise students for undergraduate thesis

Propzy Vietnam November 2021 - July 2022
AI Engineer *Ho Chi Minh City, Vietnam*

- House pricing prediction:
 - Using Graph Attention Networks (GAT) to build a house-pricing model
 - Integrating edge information into GAT
 - Applied Reinforcement Learning and Self-learning for house-pricing models
- Recommendation system:
 - Augmented realty data with sequential models
 - Investigated and adapted state-of-the-art models for the recommendation system such as Transformer, Graph Convolution Networks, GAT, and Edge-GAT for real estate domain

- Proposed new metrics for evaluating model performance
- Evaluated and analyzed model's errors

· Question-Answering System

- Investigated state-of-the-art models for the Question-Answering task
- Built QA model for Vietnamese car manual dataset
- Optimized parallel computing with multiple GPUs for speeding up a model training process
- Construct a golden QA dataset for car manual (Vietnamese)
- Fine-tuning ALBERT pre-trained model for SQUAD2.0 dataset

· Virtual Assistant

- Handled errors and edge cases
- Labeled and constructed golden data for intent and slot-filling tasks for car manual domain

PROJECTS

NMT with Consistency Regularization

December 2021 - July 2022

- A semi-supervised learning method for low-resource NMT
- Augmenting unlabeled data by back-translation, then combine with labeled data for training

PUBLICATION

My Google ScholarJournal papers:

- Long H.B. Nguyen*, Viet Pham* and Dien Dinh, Improving neural machine translation with AMR semantic graphs, Mathematical Problems in Engineering, 2021, 2021. doi: 10.1155/2021/9939389. URL <https://doi.org/10.1155/2021/9939389>
- Long Nguyen, Viet Pham, H. Minh, D. Dinh, and T. Manh, Integrating AMR semantic graphs to convolutional neural machine translation, ICIC Express Letters, vol. 12, pp. 133-141, 2021

Conference papers:

- Nguyen, B., Le, B., Nguyen, L., Pham, V., Dinh, D. (2022). Providing Syntactic Awareness to Neural Machine Translation by Graph-Based Transformer. In: Dang, N.H.T., Zhang, YD., Tavares, J.M.R.S., Chen, BH. (eds) Artificial Intelligence in Data and Big Data Processing. ICABDE 2021. Lecture Notes on Data Engineering and Communications Technologies, vol 124. Springer, Cham.
- L. H. B. Nguyen, V. Pham and D. Dinh, "Integrating AMR to Neural Machine Translation using Graph Attention Networks," 2020 7th NAFOSTED Conference on Information and Computer Science (NICS), 2020,

- Viet Pham, Long H. B. Nguyen, and Dien Dinh. 2020. Semantic Convolutional Neural Machine Translation Using AMR for English-Vietnamese. In Proceedings of the 2020 International Conference on Computer Communication and Information Systems (CCCIS 2020). Association for Computing Machinery, New York, NY, USA

E-print articles:

- Viet H. Pham, Thang M. Pham*, Giang Nguyen*, Long Nguyen, Dien Dinh (2022). Semi-supervised Neural Machine Translation with Consistency Regularization. *Under review*.

(*) Equal contributors

HONORS AND AWARDS

Excellent undergraduate thesis award	2020
Scientific Research Contest for Excellent Students - Semi-final Round	2020
Honour Certificate for Achieving Excellent Results in National High School Exam	2016
Long An Provincial Excellent Student in Mathematics - Third prize	2015

TECHNICAL STRENGTHS

Programming Language	Python, C++ (basic)
Deep Learning Frameworks	Pytorch, Tensorflow, Keras
NLP Tools	NLTK, spaCy, Pandas
Version Control	Git

LANGUAGES

Vietnamese	Native speaker
English	Fluent

REFERENCES

Ph.D. Dien Dinh: Associate Professor	ddien@fit.hcmus.edu.vn
Ph.D. Long Nguyen: Lecturer, University of Science	nhblong@fit.hcmus.edu.vn