Hieu Minh Tran

michael.tranhieu@gmail.com | hieuvodoi.github.io

RESEARCH INTERESTS

My research explores mechanisms to understand human languages for computers so that computers can perform cognitive language-related tasks for us. Among others, I am especially interested in distilling structured information and mining useful knowledge from massive human-written text of various domains. I am also interested in other language-related problems with deep learning, including reading comprehension, question answering, machine translation, and natural language generation.

EDUCATION

University of Massachusetts, Amherst (UMass) Aug. 2022 - Aug 2027
Computer Science PhD student *Massachusetts, USA*

Hanoi University of Science and Technology (HUST) Aug. 2014 - Aug. 2019
Engineer's Degree in Computer Science *Hanoi, Vietnam*
- CPA: 3.6/4.0 (top 1.3%), graduated with Excellent Degree

Bien Hoa High School for The Gifted Sep. 2011 - Sep. 2014
Specialized in Informatics *Hanam, Vietnam*

RESEARCH EXPERIENCE

AI Research Resident Nov 2019 - Jun 2022
VinAI Research (www.vinai.io) *Hanoi, Vietnam*
- Supervisor: Prof. Thien Huu Nguyen (google scholar)
- Research Topic: Information Extraction

Applied Rotation Program May 2021 - Sep 2021
VinAI Research *Hanoi, Vietnam*
- Supervisor: Dr. Duy Tin Vo (google scholar), Dr. Dat Quoc Nguyen (google scholar)
- Working on knowledge-based question answering and machine translation problem

PUBLICATIONS

A Vietnamese-English Neural Machine Translation System 2022
Thien Hai Nguyen, Tuan-Duy H. Nguyen, Duy Phung, Duy Tran-Cong Nguyen, **Hieu Minh Tran**, *Manh Luong, Tin Duy Vo, Hung Hai Bui, Dinh Phung, Dat Quoc Nguyen*
Proceedings of InterSpeech 2022 - Show & Tell Demonstrations

Exploiting Document Structures and Cluster Consistencies for Event Coreference Resolution 2021
Hieu Minh Tran, *Duy Phung, Thien Huu Nguyen*
Proceedings of ACL-IJCNLP 2021 (oral presentation)

Learning Cross-lingual Representations for Event Coreference Resolution with Multi-view Alignment and Optimal Transport 2021
Duy Phung, **Hieu Minh Tran**, *Thien Huu Nguyen*
Proceedings of the first Workshop on Multilingual Representation Learning (MRL 2021) at EMNLP 2021

Vietnamese Speech-based Question Answering over Car Manuals 2021
Tin Duy Vo, Manh Tien Luong, Duong Minh Le, **Hieu Minh Tran**, *Nhan Tri Do, Duy Nguyen, Thien Hai Nguyen, Hung Hai Bui, Dat Quoc Nguyen, Dinh Quoc Phung*
Proceedings of Data-Centric AI workshop NeurIPS 2021

The Dots Have Their Values: Exploiting the Node-Edge Connection in Graph-based Neural Models for Document-level Relation Extraction 2020
Hieu Minh Tran, *Minh Trung Nguyen, Thien Huu Nguyen*
Proceedings of EMNLP 2020 (Findings)

HONORS AND AWARDS

The Excellence Scholarship - Level A *2015*
School of Information and Communication Technology, HUST
Each semester top 5% of students with greatest academic performance are awarded

National Excellent Student Award *2014*
Vietnam Ministry of Education and Training
Third prize in Informatics subject, National Excellent High School Student Contest

PROJECTS

Vietnamese Speech-based Question Answering over Car Manuals   June. 2021 - Sep. 2021

Building a novel Vietnamese speech-based question answering system QA-Car Manual that enables users to ask car-manual-related questions (e.g. how to properly operate devices and/or utilities within a car). Given a car manual written in Vietnamese as the main knowledge base, we develop QA-CarManual as a lightweight, real-time and interactive system that integrates state-of-the-art technologies in language and speech processing to understand and interact with users via speech commands and automatically query the knowledge base and return answers in both forms of text and speech as well as visualization.

Vietnamese-English Machine Translation System   Jul. 2021 - Sep. 2021

Delivering a Speech-based Vietnamese-English machine translation system. Due to the limitations of Google Translate and other translation services in the Vietnamese language, we want to build a product that can translate properly from Vietnamese text documents to English documents and vice versa. My contributions included two main parts. First, I needed to collect and preprocess bilingual data from various sources. Secondly, after finetuning the model on preprocessed data, I had to deploy a real-time service based on the model and optimize all necessary resources such as processing time, GPU memory, etc.

ACTIVITIES

Technical Talk   *August 2021*

AI Day 2021: Empowering Innovations – website: https://www.vinai.io/aiday2021

Present at AI Day 2021 - the event which welcomed technical talks from top researchers around the world and attracted over 20,000 views online

Invited speaker   *October 2021*

Public Natural Language Processing workshop organized by VinAI Reseasrch.

TECHNICAL SKILLS

Programming: Python (proficient), Java, MATLAB (Familiar)

Libraries: Pytorch, TensorFlow, Numpy, Matplotlib, Transformers, etc.

Developer Tools: Git, Docker

CERTIFICATES

Deep Learning Specialization by deeplearning.ai

- Neural Networks and Deep Learning
- Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization
- Structuring Machine Learning Projects
- Convolutional Neural Networks
- Sequence Models

LANGUAGES

- **Vietnamese**: Native

- **English**: Fluent
  IELTS: Overall 7.0, Listening 7.5, Reading 8.5, Writing 6.5, Speaking 6.0