

Contact

www.linkedin.com/in/huy-vu-94874a130 (LinkedIn)
www.huythanhvu.com (Personal)
vuthanhvu94.wixsite.com/creative-playground (Portfolio)

Top Skills

Machine Learning
Natural Language Processing (NLP)
Data Science

Honors-Awards

National Program for the
Development of Mathematics
Exchange Program Scholarship
2014-2015
Excellence in Academic Activities
2012
6th KumHo Asiana Scholarship
Department \$5000 Merit Fellowship

Publications

3D-brain Segmentation using Deep
Neural Network and Gaussian
Mixture Model
Suicide Risk Assessment with Multi-
level Dual-Context Language and
BERT
Predicting Responses to
Psychological Questionnaires from
Participants' Social Media Posts and
Question Text Embeddings
Empirical Evaluation of Pre-trained
Transformers for Human-Level
NLP: The Role of Sample Size and
Dimensionality
Modeling Latent Dimensions of
Human Beliefs

Huy Vu

Machine Learning Engineer at Nvidia | PhD @ Stony Brook
University
San Francisco Bay Area

Summary

Having a great interest in Machine Learning, especially Natural Language Processing and its application into Psychology and Social Science, which for me is a nice intersection between understanding machine and human's mind. Having worked interdisciplinary with other natural and social scientists, I always believe that the more different scientific fields intersect, the more novel and exciting the findings will be. My main research direction is text generation conditioned on human mental health and psychological traits such as personalities, with the goal to understand human psychology with another perspective, as well as building more human-like, sympathetic conversation robots.

Check out personal site here: <https://www.huythanhvu.com/>

Experience

NVIDIA

Senior Deep Learning Engineer
July 2023 - Present (1 year 2 months)
Santa Clara, California, United States

Working on optimizing deep learning model's performances.

Stanford University

Research Associate
May 2023 - Present (1 year 4 months)
California, United States

+ Working with Computational Psychology & Well-Being Lab at Stanford, building chatbots upon LLMs (e.g., GPT-3.5) simulating human therapists to conduct major steps in CBT (Cognitive Behavioral Therapy) practice treatment. Models are constructed using a variety of methods: prompt engineering, few-shot learning, and fine-tuning through OpenAI's API.

Stony Brook University
Graduate Researcher

September 2019 - May 2023 (3 years 9 months)

Stony Brook, New York, United States

- + Working with psychology professors, social scientists from Stanford Uni., Uni. of Melbourne, and Uni. of Pennsylvania to build language generative models generating texts conditioned on Big5 personalities or mental health variables (e.g. depression, life-satisfaction). These models help understanding the language of different psychological traits as well as are potential for humanized chatbots for psychology research purposes.
- + Creating models modified based on GPT-2 simulating language of human having different psychological traits. Now in process of writing paper.

Meta

Machine Learning Engineer

May 2022 - August 2022 (4 months)

California, United States

- + Generating synthetic training data using Text-to-Speech pipeline to improve Automatic Speech Recognition (ASR) task. Main goal is to improve models' Fairness by reducing the performance gaps across different age-groups.
- + Reduced the performance gaps in ASR task across age-groups up to 13% by: (1) Reviewing literature to look for appropriate dataset to create synthetic data from. (2) Analyzing speed and pitch features distribution of real data to match with the synthetic data distribution.

NVIDIA

Deep Learning Engineer

June 2021 - August 2021 (3 months)

Santa Clara, California, United States

- + Thorough surveying of landscape and implementing strong models (e.g., DPR, Fusion-in-Decoder, Megatron-QA) for Open Domain Question Answers task.
- + Working on tuning and accelerating models (running on NVIDIA's DGX A100 GPUs and Selene supercomputer) to reduce computational costs while still obtaining high performance.

Amazon

Applied Scientist

May 2020 - August 2020 (4 months)

Seattle, Washington, United States

- + Working on distilling strong, powerful, but heavy transformers-based language model into more compact, fast-running model while still maintaining reasonable performances.
- + Proposed a novel approach to take advantage of large models' performance (RoBERTa, XLM) to be used in a real time on devices with small computing capability. Received return offer.

Brookhaven National Laboratory

Researcher

September 2019 - May 2020 (9 months)

New York, U.S.

- + Working with material scientists, implementing text mining algorithms for BNL's material science literature database.
- + Extracting unstructured text, re-organizing to structured database to improve data retrieval efficiency for material scientists.
- + Extracting contextualized embeddings of the literature, then analyzing the embedding space to suggest new materials having desired properties.

University of Pennsylvania

Researcher Internship

June 2019 - August 2019 (3 months)

Pennsylvania, U.S.

- + Working with psychologists, exploring novel ideas of applying contextualized word embeddings in psychology research.
- + Proved validity of the proposed hypotheses, with correlation up to 0.51 (significant in psychology field).

Stony Brook University

Teaching Assistant

August 2018 - April 2019 (9 months)

New York, U.S.

Teaching Assistant on courses:

- + Discrete Mathematics
- + Computer Science Foundations

University of Science Ho Chi Minh City

2 years 6 months

Teaching Assistant

February 2016 - February 2018 (2 years 1 month)

Ho Chi Minh City, Vietnam

Teaching Assistant on courses:

- + Data Structure and Algorithm
- + Introduction to Programming

Research Assistant

September 2015 - September 2017 (2 years 1 month)

Ho Chi Minh City, Vietnam

Working on projects:

- + Predicting Transmembrane Domain of Proteins using LSTM Recurrent Neural Network
- + 3D-brain Segmentation using Deep Neural Network and Gaussian Mixture Model

Omn1Solution

Business Analyst

2017 - 2018 (1 year)

Ho Chi Minh City, Vietnam

Responsibilities at Omn1Solution (Salesforce's official partner in Vietnam):

- + Assisting in analyzing customers' needs and designing system solution
- + Implementing and deploying Salesforce Sales Cloud system

Education

Stony Brook University

Doctor of Philosophy - PhD, Computer Science · (2018 - 2023)

University Of Economics Ho Chi Minh City

Certificate, Business Administration Short-term Course · (2017 - 2017)

VNUHCM - University of Science

Bachelor's degree, Mathematics and Computer Science · (2012 - 2017)

Tokyo University of Agriculture and Technology

Exchange Student and Research Internship, Computer Science · (2014 - 2015)

VNU-HCM High School for the Gifted

· (2009 - 2012)