Truong Buu Phan

Website: truongbuu.github.io - Email: phanbuu1194@gmail.com
Google Scholar - LinkedIn

RESEARCH INTERESTS

My research interests lie within information theory, probabilistic inference and representation learning, with a particular focus on developing efficient machine learning algorithms and understanding their fundamental limits. Before starting my PhD, I worked on adversarial attacks, approximate inference, out-of-distribution detection, and uncertainty quantification for neural networks.

Ressearch Topics: Information Theory, Monte-Carlo Methods, Language Models.

ACADEMIC BACKGROUND

Ph.D. Electrical and Computer Engineering 2022 - Present
University of Toronto, Canada
- Advisor: Prof. Ashish Khisti.
- Current topic: sampling, compression and representation learning.

MASc. Electrical and Computer Engineering 2017-2019
University of Waterloo, Canada
- Advisor: Prof. Krzysztof Czarnecki.
- Research in Bayesian deep learning and out-of-distribution detection.

BEng. Electrical Engineering 2012-2016
Vietnam National University, Vietnam.

RESEARCH EXPERIENCE

Meta AI (FAIR), USA Mar 2024 - Oct 2024
AI Research Intern (Language Model and Probabilistic Reasoning).
Host: Dr. Karen Ullrich.
Project: Propose and develop an exact probabilistic framework to convert tokenized LLMs to byte-level LLMs. Unlike prior works, this framework does not modify weights of the original LLMs. Applications: LLM collaborations and code completions

LG Electronics AI Lab, Canada May 2021 - May 2022
Research Engineer
Project: Automated check-out and Neural Symbolic AI.

Algolux (acquired by Torc Robotics in 2023), Canada Oct 2019 - May 2021
Research Scientist
Collaborators: Prof. Felix Heide and Dr. Fahim Mannan.
Projects: Adversarial attack and efficient deep learning for autonomous driving.

SELECTED PUBLICATIONS

Buu Phan, Marton Havasi, Matthew Muckley and Karen Ullrich. "Understanding and Mitigating Tokenization Bias in Language Models." ICML 2024 Workshop on Theoretical Foundations of Foundation Models. https://arxiv.org/abs/2406.16829

Buu Phan*, Ashish Khisti*, and Christos Louizos. "Importance matching lemma for lossy compression with side information." In International Conference on Artificial Intelligence and Statistics, 2024. https://arxiv.org/abs/2401.02609

Salehkalaibar Sadaf*, **Buu Phan***, Jun Chen, Wei Yu, and Ashish Khisti: *On the Choice of Perception Loss Function for Learned Video Compression.* NeurIPS 2023. **Spotlight in the Neural Compression Workshop, ICML 2023.** https://arxiv.org/abs/2305.19301

Buu Phan, Fahim Mannan, and Felix Heide: *Adversarial imaging pipelines.* In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 16051-16061. 2021. https://arxiv.org/abs/2102.03728

Buu Phan, Samin Khan, Rick Salay, and Krzysztof Czarnecki: *Bayesian uncertainty quantification with synthetic data.* In Computer Safety, Reliability, and Security: SAFECOMP 2019 Workshops **(Best paper award).** https://link.springer.com/chapter/10.1007/978-3-030-26250-1_31.

OTHER PUBLICATIONS

See my Google Scholar.

AWARDS

Ontario Graduate Scholarship 2023
Award for top students in Ontario, Canada.

Best paper award 2019
Received at the Workshop on Artificial Intelligence Safety Engineering for the paper "Bayesian uncertainty quantification with synthetic data."

Toyota Canada Graduate Scholarship 2018
Scholarship for graduate students working in AI Safety.

Faculty of Engineering Awards, University of Waterloo 2018, 2019
Scholarship for top-performing graduate students.

International Master's Student Awards, University of Waterloo 2017-2019
Scholarship for international graduate students.

Valedictorian, International University 2016

ACADEMIC SERVICES

- Reviewer - TMLR/AISTATS 2024/CVPR 2022/ ICML 2023 Workshops

SOFTWARE SKILLS

Programming: PyTorch, Tensorflow, Matlab, PyThon, C/C++.
System: Unix, Docker.