# Ngoc Ngo Quang Tran

(678) 871-7571 | hi@ngoc.io | ngoc.io | ngoctnq | ngoctnq | hba44u0AAAAJ

## Research Interest

My research focuses on robust and trustworthy machine learning, with an emphasis on improving the generalization and stability of large models under real-world distribution shifts. I investigate how LLM-derived representations improve representation robustness and transferability, and cross-modal integration in security-critical applications. I also investigate post-training backdoor removal techniques that enhance model reliability under data scarcity constraint. Other interests include improving training stability for sparse mixture-of-experts architectures, and safeguarding privacy against data misuse in generative foundation models.

## Education

**Vanderbilt University** *GPA: 3.94/4.00* — Nashville, TN
*Ph.D. Student in Computer Science* — *Expected May 2028*

**Rensselaer Polytechnic Institute** *GPA: 4.00/4.00* — Troy, NY
*Master of Science in Computer Science* — *May 2021*

**Wabash College** *GPA: 3.74/4.00* — Crawfordsville, IN
*Bachelor of Arts with Distinction in Mathematics* — *May 2017*
Summa Cum Laude, Phi Beta Kappa

## Recent Work Experience

**Vanderbilt University** — Nashville, TN
*Graduate Research Assistant* — *August 2023 – Present*

- Designed robust algorithms for graph-based malware classification, mitigating distribution shift using LLM embeddings and function features for compatibility with existing model-based approaches.
- Developed personally-identifiable information detection and pseudonymization pipelines, combining OCR, NER, and font-style modeling for text replacement in privacy-sensitive PDF documents.

**VinAI Research** — Hanoi, Vietnam
*Research Resident* — *August 2021 – July 2023*

- Published research on contrastive learning, multitask learning, model robustness, and data privacy in top-tier venues.
- Prototyped AI systems for smart home and autonomous driving, integrating vision models into real-world pipelines.

**Sun Asterisk Inc.** — Hanoi, Vietnam
*AI Research Team Lead* — *June 2019 – August 2021*

- Led research initiatives on publications, and spearheaded internal projects in NLP and conversational AI.
- Managed the AI Research Team, designed OKRs, and delivered technical lectures and training.

**Rensselaer Polytechnic Institute** — Troy, NY
*Graduate Research Assistant* — *June 2018 - August 2018*

- Surveyed and reviewed existing literature on matrix completion from limited observations.
- Investigated the matrix sketching conditions and proved their probabilistic estimation bounds.

## Recent Publications & Manuscripts *Asterisk (*) denotes equal contributions.*

- Anwar Said, **Ngoc N. Tran**, Yuying Zhao, Tyler Derr, Mudassir Shabbir, Waseem Abbas, Xenofon Koutsoukos. "A Survey in Graph Unlearning." *Under review.*
- **Ngoc N. Tran**, Anwar Said, Waseem Abbas, Tyler Derr, Xenofon D. Koutsoukos. "Mitigating Distribution Shift in Graph-Based Android Malware Classification via Function Metadata and LLM Embeddings." *Under review.*
- Tam Nguyen, **Ngoc N. Tran**, Khai Nguyen, Richard Baraniuk. "Improving Routing in Sparse Mixture of Experts with Graph of Tokens." *Under review.*
- Dung Thuy Nguyen, **Ngoc N. Tran**, Taylor Johnson, Kevin Leach. "PBP: Post-training Backdoor Purification for Malware Classifiers." *Network and Distributed System Security Symposium (NDSS)*, 2025.

- Hoang Phan, Lam Tran, Quyen Tran, **Ngoc N. Tran**, Tuan Truong, Qi Lei, Nhat Ho, Dinh Phung and Trung Le. "Beyond Losses Reweighting: Empowering Multi-Task Learning via the Generalization Perspective." *IEEE/CVF International Conference on Computer Vision (ICCV), 2025, Highlight.*
- **Ngoc N. Tran***, Lam Tran*, Hoang Phan, Anh Bui, Tung Pham, Toan Tran, Dinh Phung and Trung Le. "Generalization Bounds for Robust Contrastive Learning: From Theory to Practice." *arXiv Preprint, 2024.*
- Thanh Van Le*, Hao Phung*, Thuan Hoang Nguyen*, Quan Dao*, **Ngoc N. Tran** and Anh Tran. "Anti-DreamBooth: Protecting Users From Personalized Text-to-Image Synthesis." *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023.
- **Ngoc N. Tran**, Hoang Phan, Tung Pham, Dinh Phung and Trung Le. "Sharpness & Shift-Aware Self-Supervised Learning." *arXiv Preprint*, 2023.
- Hoang Phan, **Ngoc N. Tran**, Trung Le, Toan Tran, Nhat Ho and Dinh Phung. "Stochastic Multiple Target Sampling Gradient Descent." *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.
- **Ngoc N. Tran**, Anh Bui, Dinh Phung and Trung Le. "Multiple Perturbation Attack: Attack Pixelwise Under Mixed $\ell_p$-norms For Better Adversarial Performance." *arXiv Preprint*, 2022.

## Recent Machine Learning Projects

**LLM Embeddings for Graph-Based Android Malware Classification** | *PyTorch-Geometric, Transformers, ollama*
- Built model inference servers to extract embedding vectors for source codes at scale using multiple LLMs through HuggingFace's Transformers, ollama, and/or standalone implementations.
- Extracted and processed in parallel function call graphs, method signatures, other binary metadata, and decompiled Java source code from the original Android packages, which are collected from the AndroZoo repository.

**Improving Routing in Sparse Mixture-of-Experts with Graph of Tokens** | *PyTorch, JAX, Transformers*
- Implemented custom graph-based routing mechanisms for various Sparse Mixture-of-Experts (SMoE) Transformer architectures to address the issue of routing fluctuation during training.
- Evaluated model performance on large-scale benchmarks of a diverse range of tasks, with in-depth analysis of routing fluctuations across layers and data modalities throughout training.

**Multi-View Ensemble for Reliable Model Predictions** | *PyTorch, RobustBench, torchattacks*
- Developed ensemble-based methods to improve model robustness by aggregating diverse outputs across multiple views of the same data, through augmenting the original model input and their intermediate representations.
- Implemented model merging techniques such as weight averaging and ZipIt-based methods to enhance reliability of the final output and reduce prediction variance.
- Combined view- and model-level aggregation using majority voting, confidence weighting, or lightweight meta-models; and evaluated these approaches against in-distribution, out-of-distribution, and adversarially-perturbed data.

**Personally-Identifiable Information Detection and Pseudonymization** | *PyTorch, Selenium, Tesseract, Augraphy*
- Developed a modular, end-to-end personally-identifiable information (PII) pseudonymization pipeline, combining OCR, Named Entity Recognition (NER) for detection, and font recognition for text replacement generation.
- Developed a synthetic data pipeline for font family detection, including font scraping and preprocessing, random input text sampling, and realistic augmentations, to enable robust PII pseudonymization.
- Developed a web-based annotation platform for crowdsourcing bounding box annotation and text transcription labels to finetune pipeline components, compatible with the Prolific API.

## Skills

**Programming Languages**: Python, C/C++
**Frameworks/Libraries**: PyTorch, JAX, NumPy/SciPy, Transformers, matplotlib, seaborn, pandas, cvxpy
**Language Proficiency**: English (fluent), Vietnamese (native), French (elementary), German (elementary)

## Interests

Cryptography, Music Production, Basketball & Pickleball.