Truong Quang Vo

voquangtruong152000@gmail.com
(+1) 267-808-5852
tqv24.github.io

RESEARCH INTERESTS

My research interests include: (i) developing fine-grained evaluation frameworks identifying the strengths and limitations of LLMs, (ii) exploring and generating synthetic data to induce and refine desired model behaviors, and (iii) advancing efficient machine learning algorithms for large-scale model training and adaptation. Through this work, I seek to develop rigorous methodologies and scalable techniques that strengthen the reliability and trustworthiness of language models.

EDUCATION

Northwestern University, Evanston, Illinois US
M.S. in Machine Learning and Data Science, (expected graduation date: December 2025)
- GPA: 3.98/4.0

Drexel University, Philadelphia, Pennsylvania US
B.S., Chemical Engineering, (graduation date: September, 2023)
- GPA: 3.77/4.0
- Awards: *Students Tackling Advanced Research Scholars*

RESEARCH EXPERIENCE

Stanford Trustworthy AI Research Lab – Stanford University
Visiting Student Researcher   **June 2025 – Present**

Project: *Evaluating Cross-Cultural Understanding and Reasoning in Large Language Models*
- Developed a large-scale cultural reasoning benchmark (80K+ prompts) integrated into HELM Leaderboard, enabling standardized evaluation of cultural knowledge and reasoning capabilities of frontier LLMs across 25+ countries.
- Fine-tuned LLM with supervised fine-tuning (SFT) on 20K+ reasoning-answer pairs followed by reinforcement learning (PPO), improving LLM alignment on World Value Survey benchmark.

Center for Deep Learning – Northwestern University
Graduate Research Assistant   **June 2024 – May 2025**

Project: *Synthetic Clinical Data Generation for Robust and Rare-Condition ICD Coding Models*
- Built a synthetic data pipeline using medical symptoms knowledge graphs to generate targeted clinical samples for rare ICD codes, expanding coverage of underrepresented conditions.
- Fine-tuned large language models for ICD coding using supervised fine-tuning (SFT) with generated synthetic data, improving robustness and performance on rare disease categories.

Project: *Efficient Training of CNN and Transformer Ensembles via Subgrid-Importance Boosting*
- Developed a feature-prioritized boosting algorithm (Subgrid-BoostCNN) that accelerates CNN ensemble training by 40% while preserving predictive performance.
- Introduced a sub-sequence selection strategy for transformer boosting (Subsequence-BoostTransformer), improving computational efficiency and enabling scaling to longer sequences and larger ensembles.

Project: *Stochastic Large-Scale Algorithms with Distributed Features and Observations*
- Developed the Stochastic Doubly Distributed Algorithm (SODDA) to enable efficient large-scale optimization under both feature and observation partitioning in distributed environments.
- Implemented and evaluated the algorithm on multi-node Apache Spark clusters, demonstrating higher throughput in large-scale machine learning pipelines.

PUBLICATIONS

Conferences & Workshops

[1] **Truong Vo** & Sanmi Koyejo. Cultural Understanding and Reasoning Evaluation: A Framework for "Thick" Culture Alignment Evaluation in LLMs. In **AAAI 2026** *Workshop on AI for Governance.*

[2] **Truong Vo** & Sanmi Koyejo. Toward Comprehensive Cultural Alignment in LLMs: An Interdisciplinary Framework for Cultural Evaluation. In **AAAI 2026** *Workshop on AI for Governance.*

[3] Biyi Fang, **Truong Vo**, Jean Utke, & Diego Klabjan. Tricks and Plug-ins for Gradient Boosting in Image Classification. In *Proceedings of the 2025 IEEE International Conference on Big Data* (**BigData 2025**).

[4] Biyi Fang, **Truong Vo**, Kripa Rajshekhar, & Diego Klabjan. Topic Analysis with Side Information: A Neural-Augmented LDA Approach. In **ICDM 2025** *Workshop on Multimodal AI.*

[5] Biyi Fang, Diego Klabjan, & **Truong Vo**. Stochastic Large-Scale Machine Learning Algorithms with Distributed Features and Observations. In *Proceedings of the 2024 IEEE International Conference on Big Data* (**BigData 2024**).

Preprints & Under Review

[6] **Truong Vo**, Weiyi Wu, & Kaize Ding. Synthetic Clinical Data Generation for Training Robust ICD Coding Models. *Under review at the 64th Annual Meeting of the Association for Computational Linguistics* (**ACL 2026**).

[7] Biyi Fang, **Truong Vo**, Jean Utke, & Diego Klabjan. Tricks and Plug-ins for Gradient Boosting for Transformers. *Under review at the 27th International Conference on Pattern Recognition* (**ICPR 2026**).

PROFESSIONAL EXPERIENCE	**Sabre Holdings** – Remote

AI Engineer Intern   **September 2025 – Present**

- Developed multi-agent AI system with AutoGen to power a personalized air travel recommender, combining Sabre's search engine with purchase prediction APIs to deliver optimal itineraries.
- Integrated the SendGrid API for automated communication, creating an end-to-end pipeline that sends flight recommendations directly to users.

Mintel Group – Remote

Applied AI Intern   **September 2024 – May 2025**

- Led the development of an AI workflow using LangGraph, enabling clients to perform deep analysis on the global F&B market and achieving a 75% reduction in market analytics time.
- Built an agentic retrieval-augmented generation (RAG) system with the Pinecone vector database to handle human interaction and extract key information from user queries.
- Developed an analytic agent for SQL querying and trend detection across 20 GB of GNPD data, automating market analytics and insights generation.

TEACHING EXPERIENCE	**Computational Method for Engineering** – Drexel University

Teaching Assistant   **September 2022 – December 2022**

- Mentored students in using Python programming to solve differential equations, providing guidance and detailed feedback on student's homework.
- Prepared workshop materials and held weekly office hours to reinforce in-class concepts and clarify key mathematical principles.

Introductory Programming for Engineers – Drexel University

Teaching Assistant   **January 2022 – April 2022**

- Mentored students in core Python programming concepts and provided debugging assistance and structured feedback.
- Created coding demonstrations and guided students through step-by-step implementation examples.

REFERENCES

Assistant Professor Sanmi Koyejo
Department of Computer Science – Stanford University
Email: sanmi@cs.stanford.edu

Professor Diego Klabjan
Department of Industrial Engineering and Management Sciences – Northwestern University
Email: d-klabjan@northwestern.edu

Professor Yuri Balasanov
Department of Industrial Engineering and Management Sciences – Northwestern University
Email: yuri.balasanov@northwestern.edu