

EDUCATION **Stanford University**

Ph.D. '27, Computer Science, M.S. '24, Computer Science, M.S. '25, Statistics

Advisors: Sanmi Koyejo, Nick Haber, Susan Athey

Affiliations: SAIL, StatsML, GSE, GSB

DePauw University

B.A. '21, Computer Science, Economics, Computational Chemistry

Advisors: Jeff Gropp, Humberto Barreto, Todd Neller

Affiliations: Management Science Fellows

RESEARCH I develop trustworthy autonomous artificial agents that make efficient decisions in various real-world applications. My research encompasses multiple subjects, including active learning, reinforcement learning, model evaluation, and natural language generation. I applied these techniques to solve problems in human-centered applications, such as education, healthcare, and scientific discovery.

Conference & Journal Publications

* denotes equal contribution

- [10] **Truong** et al. Crossing Linguistic Horizons: A Comprehensive Evaluation of Vietnamese Large Language Models. *NAACL*'24.
- [9] Wang, ..., **Truong** et al. Decoding Trust: A Comprehensive Assessment of Trustworthiness in GPT Models. *NeurIPS*'23. **Best Paper Award**.
- [8] Griffiths*, ..., **Truong*** et al. GAUCHE: A Library for Gaussian Processes in Chemistry. *NeurIPS*'23.
- [7] Pham, ..., **Truong** et al. Ag/ZnO Heterojunction Decorated on Polypropylene as a membrane for Photocatalytic NO Removal. *Industrial & Engineering Chemistry Research*'22.
- [6] **Truong**, Neller. A Data-Driven Approach for Gin Rummy Hand Evaluation. *AAAI*'21.
- [5] **Truong**, Barreto. Teaching Income Inequality with Data Driven Visualization. *The American Economist*'22.
- [4] Pham, ..., **Truong** et al. One-pot hydrothermal synthesis of Si/TiO₂ from commercial material sources for visible light-driven photocatalytic activity. *Material Research Express*'19.
- [3] Pham, **Truong** et al. Synthesis of Ag/TiO₂ nanocomposite by γ irradiation for enhanced photocatalytic activity under sunlight. *Nuclear Instruments & Methods in Physics Research*'18.
- [2] Pham, **Truong** et al. An improved green synthesis method and Escherichia coli antibacterial activity of silver nanoparticles. *Journal of Photochemistry & Photobiology*'18.
- [1] Pham, ..., **Truong** et al. Ag/TiO₂ with high photocatalytic and antibacterial activity synthesized by photoreduction method. *Journal of Photochemistry & Photobiology*'18.

Workshop Papers & Technical Reports

- [10] Niknazar*, Haley*, Ramanan*, **Truong***, Shrinivasan* et al. Building a Production-Grade Appropriateness Model for Educational AI. *Submitted at COLM*'24
- [9] **Truong** et al. Scalable Non-myopic Bayesian Optimization for Dynamic Cost Settings. *Submitted at Advanced in Approximated Bayesian Inference* '24
- [8] Chang, ..., **Truong** et al. Red Teaming Large Language Models in Medicine: Real-World Insights on Model Behavior. *Submitted at The New England Journal of Medicine AI* '24
- [7] Azem*, **Truong*** et al. Pretraining Probabilistic Models for Scalable Precision Agriculture *ICLR Data-centric Machine Learning Research* '24, submitted at *Advanced in Approximated Bayesian Inference* '24
- [6] Schaeffer, ..., **Truong** et al. Bridging Associative Memory and Probabilistic Modeling. *Submitted at ICML*'24.
- [5] Bhatt, ..., **Truong** et al. An Experimental Design Framework for Label-Efficient Supervised Finetuning of Large Language Models. *Submitted at ACL* '24.
- [4] Truong, ..., **Truong**. Hybrid Transformer and Holt-Winter's Method for Time Series Forecasting. *ICLR Time Series for Health* '24.
- [3] **Truong** et al. Thomas: Learning to Explore Human Preference via Probabilistic Reward Model. *ICML The Many Facets of Preference-Based Learning* '23.
- [2] Duong, **Truong** et al. Quantum Neural Architecture Search with Quantum Circuits Metric and Bayesian Optimization. *ICML Active Learning & Experimental Design* '22.
- [1] Voegelé, ..., **Truong** et al. Systematic Comparison of Biomolecular Structural Ensembles via Mutual Information. *EMBO Advances & Challenges in Biomolecular Simulations* '21. **Best Presentation Award**.

Selected Invited Talks

- [3] Crossing Linguistic Horizons: A Comprehensive Evaluation of Vietnamese LLMs, FPT AI Research, 2024
- [2] Scalable Non-myopic Bayesian Optimization for Dynamic Cost Settings, Huawei London, 2023
- [1] Probabilistic Optimal Survey Design, Human-center Artificial Intelligence, Stanford University, 2022

AWARDS

- [10] Microsoft Accelerate Foundation Models Research Grant '23 (\$80,000, 1 year)
- [9] Stanford Human-center Artificial Intelligence Graduate Fellowship '22 [Press]
- [8] Stanford School of Engineering Fellowship '21
- [7] The Ferid Murad Medal '21 [Press]
- [6] Robert Thomas Award in Computer Science '21

- [5] Bruce Long, Frank Carlton, and Randal Wilson Awards in Economics '19 '20 '21
- [4] John Ricketts Award in Chemistry '21
- [3] American Chemical Society Award '18
- [2] Silver Medal, Biochemistry, Intel Science and Engineering Fair '16
- [1] Bronze Medals, Chemistry, National Olympiad '15 '16

EXPERIENCE

- [5] Merlyn Mind – *Research Intern '24, working on efficient fine-tune and evaluate safety large language models*
- [4] Google \cap The R Project – *Software Engineer Intern '21, developed hyperSpec software for spectral analysis*
- [3] University of Oxford – *Visiting Scholar '21, studied structured probabilistic deep learning*
- [2] Community Health Network – *Data Science Intern '20, worked on data-driven decision support*
- [1] Cummins – *Data Analytics Intern '19, worked on data-driven decision support*

SERVICES & LEADERSHIP

- [6] NeurIPS '22 '23, ICLR '23, ICML '22, AAAI '22 – *Reviewer*
- [5] Data-centric Machine Learning Research: Datasets for Foundation Models, ICML'24 – *Organizer*
- [4] Data-centric Machine Learning Research: Harnessing Momentum for Science, ICLR'24 – *Organizer*
- [3] Stanford Generative AI for the Future of Learning '23 – *Grant Reviewer*
- [2] Molecular Machine Learning Conference '23 – *Organizer*
- [1] Stanford Computing & Society '22 '24 – *Organizer*

TEACHING & ADVISING

Stanford CS329H: Machine Learning from Human Preferences – *Course Assistant Fall'23, developed the initial version of the course with an open-access textbook*

Research Advisees

- [3] Duc Nguyen, M.S. student in Computer Science, HCM University of Technology
- [2] Muhideen Mustapha, Stanford LINXS/INSPIRE-CS scholar in Computer Science
- [1] Hieu Tran, Ph.D. student in Computer Science, Purdue University