



## Sang Truong

Ph.D. Student in Computer Science, admitted Autumn 2021

 Curriculum Vitae available Online

### Publications

#### PUBLICATIONS

- **Red teaming ChatGPT in medicine to yield real-world insights on model behavior.** *NPJ digital medicine*  
Chang, C. T., Farah, H., Gui, H., Rezaei, S. J., Bou-Khalil, C., Park, Y. J., Swaminathan, A., Omiye, J. A., Kolluri, A., Chaurasia, A., Lozano, A., Heiman, A., Jia, et al  
2025; 8 (1): 149
- **Systematic analysis of biomolecular conformational ensembles with PENSEA.** *The Journal of chemical physics*  
Vogele, M., Thomson, N. J., Truong, S. T., McAvity, J., Zachariae, U., Dror, R. O.  
2025; 162 (1)
- **Bayesian Optimization for Crop Genetics with Scalable Probabilistic Models**  
Azam, R., Truong, S. T., Fernandes, S. B., Leakey, A. D. B., Lipka, A., El-Kebir, M., Koyejo, S.  
edited by Antoran, J., Naesseth, C. A.  
JMLR-JOURNAL MACHINE LEARNING RESEARCH.2024: 30-44
- **Crossing Linguistic Horizons: Finetuning and Comprehensive Evaluation of Vietnamese Large Language Models**  
Truong, S. T., Nguyen, D. Q., Toan Nguyen, Le, D. D., Truong, N. N., Tho Quan, Koyejo, S.  
edited by Duh, K., Gomez, H., Bethard, S.  
ASSOC COMPUTATIONAL LINGUISTICS-ACL.2024: 2849-2900
- **An Experimental Design Framework for Label-Efficient Supervised Finetuning of Large Language Models**  
Bhatt, G., Chen, Y., Das, A. M., Zhang, J., Truong, S. T., Mussmann, S., Zhu, Y., Bilmes, J., Du, S. S., Jamieson, K., Ash, J. T., Nowak, R. D.  
edited by Martins, A., Srikumar, Ku, L. W.  
ASSOC COMPUTATIONAL LINGUISTICS-ACL.2024: 6549-6560
- **DECODINGTRUST: A Comprehensive Assessment of Trustworthiness in GPT Models**  
Wang, B., Chen, W., Pei, H., Xie, C., Kang, M., Zhang, C., Xu, C., Xiong, Z., Dutta, R., Schaeffer, R., Truong, S. T., Arora, S., Mazeika, et al  
edited by Oh, A., Neumann, T., Globerson, A., Saenko, K., Hardt, M., Levine, S.  
NEURAL INFORMATION PROCESSING SYSTEMS (NIPS).2023
- **GAUCHE: A Library for Gaussian Processes in Chemistry**  
Griffiths, R., Klarner, L., Moss, H., Ravuri, A., Truong, S., Stanton, S., Tom, G., Rankovic, B., Du, Y., Jamasb, A., Deshwal, A., Schwartz, J., Tripp, et al  
edited by Oh, A., Neumann, T., Globerson, A., Saenko, K., Hardt, M., Levine, S.  
NEURAL INFORMATION PROCESSING SYSTEMS (NIPS).2023