Dang **Nguyen**

(+1) 310-254-4895 | nguyentuanhaidang@gmail.com | hsgser.github.io | hsgser | dang-nguyen-50b7a7a0 | Dang Nguyen

Research interests

My research is centered on developing efficient and robust machine-learning algorithms for large-scale datasets and architectures. Specifically, I am focused on improving the training efficiency of large language models and enhancing their robustness. Additionally, I am interested in exploring Generative AI and Multimodal Learning, with a particular emphasis on improving data quality.

Education

University of California, Los Angeles — *California, USA*

Ph.D. in Computer Science — *Sep. 2023 - Present*

- Advised by Professor Baharan Mirzasoleiman
- UCLA Graduate Dean's Scholar Award

Toyo University — *Tokyo, Japan*

B.S. in Information Networking for Innovation and Design — *Apr. 2017 - Mar. 2021*

- Toyo Top Global Scholarship A
- GPA: 4.27/4.3, Top 1/300 in the faculty

Experience

Google Research — *California, USA*

Student Researcher — *Sep. 2024 - Dec. 2024*

- Topic: Synthetic data generation for LLMs

Cisco — *California, USA*

PhD Research Intern — *Jun. 2024 - Sep. 2024*

- Supervisor: Dr. Ali Payani
- Topic: LLM Hallucination

VinAI — *Hanoi, Vietnam*

AI Resident — *Oct. 2020 - Aug. 2023*

- Supervisor: Professor Nhat Ho (UT Austin)
- Topics: Optimal Transport and Model Merging

Publications

(*) denotes equal contribution

1. **D. Nguyen**, P. Haddad, E. Gan, and B. Mirzasoleiman. Make the Most of Your Data: Changing the Training Data Distribution to Improve In-distribution Generalization Performance. *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.
2. Y. Xue, J. Siddharth, **D. Nguyen**, and B. Mirzasoleiman. Understanding the Robustness of Multi-modal Contrastive Learning to Distribution Shift. *International Conference on Learning Representations (ICLR)*, 2024.
3. K. Nguyen*, **D. Nguyen***, N. Ho. Self-Attention Amortized Distributional Projection Optimization for Sliced Wasserstein Point-Cloud Reconstruction. *International Conference on Machine Learning (ICML)*, 2023.
4. **D. Nguyen**, T. Nguyen, K. Nguyen, D. Phung, H. Bui, and N. Ho. On cross-layer alignment for model fusion of heterogeneous neural networks. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2023.
5. K. Nguyen*, **D. Nguyen***, T. A. V. Le, T. Pham, and N. Ho. Improving mini-batch optimal transport via partial transportation. *International Conference on Machine Learning (ICML)*, 2022.
6. K. Nguyen, **D. Nguyen**, Q. Nguyen, T. Pham, H. Bui, D. Phung, T. Le, and N. Ho. On transportation of mini-batches: A hierarchical approach. *International Conference on Machine Learning (ICML)*, 2022.

Preprints

1. **D. Nguyen**, W. Yang, R. Anand, Y. Yang and B. Mirzasoleiman. Mini-batch Coresets for Memory-efficient Training of Large Language Models. 2024.

Professional services

- Reviewer at NeurIPS (2022-2024), ICML (2023-2024), ICLR (2024-2025), AISTATS (2023-2025), AAAI 2025, AdvML@NeurIPS2024

Honors & Awards

2023	**UCLA Graduate Dean's Scholar Award**, UCLA	*California, USA*
2017	**Toyo Top Global Scholarship A**, Toyo University	*Tokyo, Japan*
2015	**Silver medal**, 56th International Mathematical Olympiad	*Chiang Mai, Thailand*
2015	**First Prize**, Vietnam Mathematical Olympiad	*Hanoi, Vietnam*