

Anshul Nasery

PhD Student, University of Washington

@ anshulnasery@gmail.com  Homepage  Github  Google Scholar

Research Interests

I am interested in fundamental deep learning research for training secure, efficient and deployment ready models.

Education

Ongoing Sept 2023	University of Washington PhD in Computer Science Advisor: Prof. Sewoong Oh	
Aug 2021 Jul 2017	Indian Institute of Technology, Bombay Bachelor of Technology (With Honors) in Computer Science and Engineering, Minor in Statistics Advisor: Prof. Sunita Sarawagi	GPA: 9.58/10

Industrial Research Experience

Jul 2021 Jul 2023	Google Research Pre-Doctoral Researcher Advisor: Dr. Prateek Jain, Dr. Praneeth Netrapalli Worked on research problems around inference efficient and generalizable neural networks.	Bangalore, India
Apr 2020 Jul 2020	Adobe Research Research Intern Advisor: Dr. Balaji Vasani Srinivasan Worked on a research problem around Multi-Modal Question Answering [NAACL'21, US Patent].	Bangalore, India

Selected Publications and Preprints

- [1] **PLeas - Merging Models with Permutations and Least Squares** 
Anshul Nasery*, Jonathan Hayase*, Pang Wei Koh, Sewoong Oh
Under Review [Preprint]
- [2] **Peekaboo: Interactive Video Generation via Masked-Diffusion** 
Anshul Nasery*, Yash Jain*, Vibhav Vineet, Harkirat Behl
IEEE / CVF Computer Vision and Pattern Recognition Conference, 2024 [CVPR'24]
- [3] **Learning an Invertible Mapping can Mitigate Simplicity Bias** 
Sravanti Addepalli*, Anshul Nasery*, R Venkatesh Babu, Praneeth Netrapalli, Prateek Jain
International Conference on Learning Representations, 2023 [ICLR'23]
- [4] **Training for the Future: A Simple Gradient Interpolation loss to Generalize Along Time** 
Anshul Nasery*, Soumyadeep Thakur*, Vihari Piratla, Abir De, Sunita Sarawagi
34th Conference on Advances in Neural Information Processing Systems [NeurIPS '21]
- [5] **What if Neural Networks had SVDs?** 
Alexander Mathiasen, Frederik Hvilshøj, Jakob Rødsgaard Jørgensen, Anshul Nasery, Davide Mottin
Spotlight at 33rd Conference on Advances in Neural Information Processing Systems [NeurIPS'20]
- [6] **CogCNN: Mimicking Human Cognition to resolve Texture Shape Bias** 
Satyam Mohla*, Anshul Nasery*, Biplab Banerjee
2022 IEEE International Conference on Acoustics, Speech and Signal Processing [ICASSP'22]
- [7] **MIMOQA: Multimodal Input Multimodal Output Question Answering** 
Hrituraj Singh, Anshul Nasery*, Denil Mehta*, Jatin Lamba, Aishwarya Agarwal, Balaji Vasani
2021 Conference of the North American Chapter of the Association for Computational Linguistics [NAACL'21]
- [8] **End-to-End Neural Network Compression via $\frac{\ell_1}{\ell_2}$ Regularized Latency Surrogates** 
Anshul Nasery, Hardik Shah, Arun Suggala, Prateek Jain
Mobile AI Workshop at CVPR 2024 [CVPR-W'24]
- [9] **DAFT: Distilling Adversarially Finetuned Teachers for better OOD generalization** 
Anshul Nasery, Sravanti Addepalli, Praneeth Netrapalli, Prateek Jain
Principles of Distribution Shifts Workshop, ICML 2022 [ICML-W'22]

Selected Research Projects

Robust and Scalable Fingerprinting for LLMs

May'24 - Present

Advisors: *Prof. Sewoong Oh*

- › Investigating strategies for inserting fingerprints into pretrained LLMs with minimal performance degradation.
- › Proposed a method to improve fingerprint persistence with resistance to finetuning, merging, quantization attacks.

Out-of-Domain Robustness of Neural Nets

Sept'21 - June'23

Advisors: *Dr. Prateek Jain, Dr. Praneeth Netrapalli*

- › Developed a novel feature reconstruction regularizer to alleviate simplicity bias and improve OOD generalization.
- › Obtained upto **1% gain** in accuracy over state-of-the-art methods on the DomainBed benchmark. [ICLR'23]
- › Combined adversarial fine-tuning and knowledge distillation to boost the OOD robustness of small models. [ICML-W'22]
- › Using the proposed technique, a **ResNet-50 can outperform a ResNet-101 by 2.5%** on the DomainBed benchmark.

Model merging for efficient ensembles

Nov'23 - May'24

Advisors: *Prof. Sewoong Oh, Prof Pang Wei Koh*

- › Extended Git Re-Basin to enable partial merging, producing merged models of varying sizes.
- › Proposed a novel feature adjustment step to distill knowledge into a merged model with minimal data requirements.
- › Demonstrated empirical gains of upto 13% over state of the art merging methods for various tasks.

Inference Efficient ML Models

Jul'21 - Jul'23

Advisors: *Dr. Prateek Jain, Dr. Praneeth Netrapalli, Dr. Gaurav Aggarwal*

- › **NAS**. Achieved **0.8% gain in ImageNet accuracy** for no extra FLOPs on MobileNetV3 using a novel FLOPs regularizer.
- › **Conditional Computation**. Obtained **1% gain in ImageNet accuracy** for MobileNetV2 by introducing decision trees to route examples. Introduced a skip-and-branch architecture for **25% savings in amortized FLOPs** with MobileNetV3.
- › **Compressing LLMs**. Adapting algorithms from the compressed sensing literature to prune weight matrices of large language models by over 50%, resulting in **latency reduction of 30%**.

Academic Achievements

- › Awarded Institute Academic Prize for exceptional academic performance (top 10% of class) in IIT Bombay in 2017-2018.
- › Ranked **137** in 110000 candidates in JEE Advanced 2017 and 265 in 1.5 million candidates JEE Mains 2017.
- › Placed among the **top 35 students** in Indian National Astronomy Olympiad 2017 and qualified for Indian National Olympiad of Informatics, Indian National Physics Olympiad & Indian National Chemistry Olympiad 2017.

Other Projects

Training For the Future

Jul'20 - Jul'21

Advisor: *Prof. Sunita Sarawagi*

- › Proposed a gradient based technique for better domain generalization on temporally varying data. [NeurIPS'21]

Controllable Video Generation

Oct '23-Dec '23

Advisors: *Vibhav Vineet, Harkirat Behl*

- › Proposed a training-free method for spatio-temporally controlling the outputs of any video generation model [CVPR'24].

Better parsing with background knowledge

Sept '20 - Jan '21

Advisor: *Prof Ganesh Ramakrishnan*

- › Improved F-1 score by **1% on constituency parsing** for WSJ dataset by regularizing with linguistic rules [ACL'21].

Key Courses Undertaken

Machine Learning Theoretical ML, Advanced ML, Natural Language Processing, Intelligent Learning Agents

Math And Stats Linear Algebra, Statistical Inference, Probability and Measure Theory, Regression Analysis

Miscellaneous

- › Reviewer for ICML (2022), NeurIPS (2022-2024), ICLR (2022,2023), CVPR (2023,2024), COLM 2024.
- › Teaching assistant for undergraduate courses on Artificial Intelligence and Machine Learning and Quantum Mechanics at IIT Bombay
- › Competed and won various national quizzes, and received recognition from IIT Bombay for these.
- › As hobby projects, built a bot to play word-games over messaging apps including Discord.