Anshul **Nasery**

Pre-Doctoral Researcher, Google Research

Homepage @ anshulnasery@gmail.com Github Google Scholar

Education

Aug 2021 | **Indian Institute of Technology, Bombay** **GPA: 9.58/10**
Jul 2017 | Bachelor of Technology in Computer Science and Engineering, Minor in Statistics
Thesis Advisor: Prof. Sunita Sarawagi

Publications and Patents

C=Conference Paper, W=Workshop Paper, P=Patent

[W.2] **Learning an Invertible Mapping can Mitigate Simplicity Bias** [🔗]
Sravanti Addepalli*, **Anshul Nasery***, R Venkatesh Babu, Praneeth Netrapalli, Prateek Jain
DistShift Workshop, NeurIPS 2022, Under Submission at ICLR'22 [**NeurIPS-W'22**]

[W.1] **DAFT: Distilling Adversarially Finetuned Teachers for better OOD generalization** [🔗]
Anshul Nasery, Sravanti Addepalli, Praneeth Netrapalli, Prateek Jain
Principles of Distribution Shifts Workshop, ICML 2022, Under Submission at TMLR [**ICML-W'22**]

[C.5] **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**]

[C.4] **What if Neural Networks had SVDs?** [🔗]
Alexander Mathiasen, Frederik Hvilshøj, Jakob Rødsgaard Jørgensen, **Anshul Nasery**, Davide Mottin
33rd Conference on Advances in Neural Information Processing Systems [**NeurIPS'20**]

[C.3] **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**]

[C.2] **MIMOQA: Multimodal Input Multimodal Output Question Answering** [🔗]
Hrituraj Singh, **Anshul Nasery***, Denil Mehta*, Jatin Lamba, Aishwarya Agarwal, Balaji Vasan
2021 Conference of the North American Chapter of the Association for Computational Linguistics [**NAACL'21**]

[C.1] **Rule Augmented Unsupervised Constituency Parsing** [🔗]
Anshul Nasery*, Atul Sahay*, Ayush Maheshwari, Ganesh Ramakrishnan, Rishabh Iyer
Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021 [**Findings of ACL'21**]

[P.1] **Modality Agnostic Information Retrieval from Documents** [🔗]
Hrituraj Singh, Jatin Lamba, Denil Mehta, Balaji Vasan Srinivasan, **Anshul Nasery**, Aishwarya Agarwal
US Patent App 17153130

Selected Research Projects

Out-of-Domain Robustness of Neural Nets Sept'21 - Present
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. [NeurIPS-W'22]
- Combined adversarial fine-tuning and knowledge distillation to boost the OOD robustness of small models. [hyperref[W1]ICML-W'22]
- Using the proposed technique, a **ResNet-34 can outperform a ResNet-152 by 0.3%** on the DomainBed benchmark.

Training For the Future Jul'20 - Jul'21
Advisor: Prof. Sunita Sarawagi

- Investigated gradient based techniques for better **domain generalization on temporally varying** data. [NeurIPS'21]
- Achieved over **10% relative improvements** over baselines on 5 real world datasets including M5 and HousePrice.
- Proved upper bounds on generalization error of the proposed method for regression tasks with temporal drift.

What If Neural Networks had SVDs? Nov'19 - May'20
Advisor: Prof. Davide Mottin

- Developed a parallelizable algorithm for matrix multiplication via Householder decompositions of orthogonal matrices.
- Achieved a **29x speedup** over prior work by implementing the algorithm in CUDA for running on GPUs.
- Work presented as a **Spotlight paper (top 3% of all submissions)** at NeurIPS 2020 [C.3].

Inference Efficient ML Models Jul'21 - Present

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.** Adapted compressed sensing algorithms to prune weight matrices of large language models by over 50%, resulting in **latency reduction of 30%** with minimal performance drop.

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 Infomatics, Indian National Physics Olympiad & Indian National Chemistry Olympiad 2017.

Experience

Jul 2021 – Present | **Google Research** | **Bangalore, India**
Pre-Doctoral Researcher | Advisor: Dr. Prateek Jain, Dr. Praneeth Netrapalli
Worked on research problems around inference efficient and generalizable neural networks.

Apr 2020 – Jul 2020 | **Adobe Research** | **Bangalore, India**
Research Intern | Advisor: Dr. Balaji Vasan Srinivasan
Worked on a research problem around Multi-Modal Question Answering [NAACL'21, US Patent].

May 2019 – Jan 2020 | **Aarhus University** | **Aarhus, Denmark**
Research Intern | Advisor: Prof. Davide Mottin
Worked on research problems around generative modelling and efficient matrix multiplication.

Jul 2021 – Dec 2021 | **IIT Bombay** | **Mumbai, India**
Teaching Assistant| Artificial Intelligence and Machine Learning
Tutored, evaluated and prepared coding and theoretical assignments for a batch of 130 students.

Dec 2018 – Jan 2019 | **Tvarit Solutions Gmbh** | **Mumbai, India**
Machine Learning Engineer Intern
Implemented and tested a production machine learning pipeline for fault detection in industrial processes.

Other Projects

Generative Modelling using Invertible Neural Networks Summer 2019

Advisor: Prof Davide Mottin

- Formulated a novel algorithm to **compute Wasserstein Distance** between distributions for generative modelling. Implemented constant memory backprop and variational dequantization to train very deep networks efficiently.

Better parsing with background knowledge Fall 2020

Advisor: Prof Ganesh Ramakrishnan

- Improved F-1 score by **1% on constituency parsing** for WSJ dataset by regularizing model with linguistic rules [C.1].

CognitiveCNN: Mimicking Human Cognitive Models to resolve Texture-Shape Bias Fall 2019

Advisor: Prof Biplab Banerjee

- Quantified the **shape-texture bias** of neural networks using techniques from cognition and image processing. Achieved a **5 % gain** in accuracy under miscue on Office-31 dataset using a novel attention matching based regularizer [C.3].

Theoretical Foundations of Transfer Learning Spring 2021

Advisor: Prof Nutan Limaye

- Formulated transfer learning as a three-player adversarial game to obtain bounds in the universal learning framework.

Skills

Languages Python, C/C++, MySQL, MATLAB, PHP, Bash, HTML, CSS, JavaScript
Frameworks Tensorflow, PyTorch, Keras, Jax, CUDA, Git

Miscellaneous

- Competed and won various national quizzes, and recieved recognition from IIT Bombay for these.
- As hobby projects, built a bot to play word-games over messaging apps including Discord.