Anurag Dutt

Department of Computer Science Stony Brook University, NY, USA, 11790 $\Box +1$ (347) 740-5056

EDUCATION	
Stony Brook University <i>PhD. in Computer Science</i> GPA: 4.0/4.0 Advisors: Dr. Anshul Gandhi	2022 - present
Stony Brook University Masters in Applied Mathematics & Statistics with m GPA: 3.85/4.0 Advisors: Dr. Anshul Gandhi, Dr. C	ninor in Computer Science Phristian Luhmann & Dr. Fusheng Wang
Columbia University Masters in Operations Research & Financial Engine GPA: 3.3/4.0	2017 - 2018 cering
PUBLICATIONS	
Preprints and Submissions	* - equal contribution, $\#$ - contribution credited as research assistant
1. GAMMA: Graph Neural Network-Based Gagan Somashekar, Anurag Dutt, Mainak A accepted in WWW, 2024 (paper selected for or	H Multi-Bottleneck Localization for Microservices Applications Adak, Tania Lorido Botran, Anshul Gandhi <i>ral presentation in top 5% papers)</i>
 Efficiently sharing the GPU for Deep Lee Ubaid Ullah Hafeez, Anurag Dutt, Mainak A submitted in EuroSys, 2025 	e arning Workloads with Herald Adak, Anshul Gandhi
3. Social Learning via Bayesian Inverse Re Alexandra Ortmann, Anurag Dutt, Christian submitted in CogSci, 2024	inforcement Learning: Learning from and about a Learner n Luhmann
4. Distinguishing Mechanisms of Social Inf Anurag Dutt, Christian Luhmann submitted in Nature Human Psychology, 2024	luence to optimize Network Interventions
Conference Publications	
1. Evaluating the energy impact of device p Anurag Dutt*, Sri Pramodh Rachuri*, Ashl 14th International green and sustainable comp	parameters for DNN inference on edge ey Lobo, Nazeer Shaik, Anshul Gandhi, Zhenhua Liu uting conference (IGSCC), 2023

- 2. B-MEG: Bottlenecked-Microservices Extraction Using Graph Neural Networks. Gagan Somashekar, Anurag Dutt, Rohith Vaddavalli, Sai Bhargav Varanasi, Anshul Gandhi International Conference on Performance Engineering (ICPE), 2022
- 3. SMOOTH-GAN: Towards Sharp and Smooth Synthetic EHR Data Generation. Sina Rashidian, Fusheng Wang, Richard Moffitt, Victor Garcia, Anurag Dutt, Wei Chang, Vishwam Pandya, Janos Hajagos, Mary Saltz, Joel Saltz Artificial Intelligence in Medicine (AIME), 2020

Publications credited as Research Assistant

- 1. Stock market trading in the aftermath of an accounting scandal Renuka Sane, Anurag $Dutt^{\#}$ Emerging Markets Review, 2019
- 2. From participation to repurchase Low-income households and micro-insurance Renuka Sane, Susan Thomas, Anurag Dutt[#] Journal of Risk and Finance, 2019

Theses

1. Evaluating the energy impact of device and workload parameters for DNN inference on edge. Anurag Dutt, Anshul Gandhi

Graduate Thesis towards completion of PhD Qualifier, Department of Computer Science, SBU, 2023

Selected Research Projects _

OptiSplit: Maximizing LLM Inference Throughput via Efficient Attention Head Distribution on Heterogeneous GPUs 2023 - present

Advisors: Dr. Anshul Gandhi, Dr. Zhenhua Liu (Stony Brook University) This project investigates optimal distribution of attention heads across heterogenous GPU configurations to maximize inference speed for Large Language Models (LLMs).

IrEne: Interpretable Energy Prediction for Transformers Advisors: Dr. Anshul Gandhi, Dr. Niranjan Balasubramaniam, Dr. Aruna Balasubramaniam (Stony Brook University) The project focuses on building prediction models for energy consumption of Large Language models (LLMs) for distributed setups for tensor and pipeline parallelism. The project was recently selected for an OVPR seed grant.

Selected Awards and Honors

• Student travel grant to attend USENIX ATC '2023	2023
• Student travel grant to attend IGSCC '2023	2023
• Certificate of merit for finishing 3rd in AMS graduate program	2021
• Chairman's fellow in Department of Computer Science, SBU for 2 years running	2022 - 2023
• Awarded NTSE Scholarship from Government of India.	2013

TALKS _____

• Energy Efficient Computing for Edge devices	
– 14th International Green and Sustainable Computing Conference (IGSCC)	Oct 2020, Toronto, CA

Work Experience

Roc	Capital	LLC	
-----	---------	-----	--

Sr. Data Scientist

Stack: Python, PyTorch, MLflow, Weights & Biases, SigOpt, ONNX, streamlit, PostgreSQL, docker, Kubernetes & Kube-Flow, Prometheus, Grafana, Sagemaker

• My role involved designing, implementing, and maintaining scalable infrastructure and automation pipelines to streamline the deployment, monitoring, and management of deep learning models in production environments, for a hedge fund

Thomson Reuters

Data Science Intern Stack: Python, Sci-kit learn, PyTorch, Jenkins, Prometheus, fiddler, TensorRT

• Led the development and deployment of advanced deep learning-based pricing models for exotic credit default swaps, implementing robust infrastructure and automation pipelines to ensure seamless model deployment and real-time monitoring

PROFESSIONAL RESPONSIBILITIES ____

- Reviewing: USENIX ATC '24, ACM Signetrics (2022 2023), Middleware (2023)
- Mentorship
 - Students
 - * Nazeer Shaik
 - Master's Advanced Project, Stony Brook University, 2022-2023 (w Dr. Anshul Gandhi)
 - * Sarath Shiva Kondeti Master's Advanced Project, Stony Brook University, 2022-2023 (w Dr. Anshul Gandhi)

2018

2019, 2021-2022

2024 - present

* Mainak Adak	
Master's Advanced Project, Stony Brook University, 2022-2024 (w \setminus Dr. Anshul Gandhi)	
• Graduate Teaching Assistantship - Department of Computer Science, Stony Brook University	
– Natural Language Processing - Prof. Niranjan Balasubramanian	Fall 2022
– Probability & Statistics for Computer Science - Prof. Anshul Gandhi	Spring 2023
• Secretary - SIAM student chapter - Applied Mathematics & Statistics	2019 - 2021