

Objective: *A PhD candidate with extensive expertise in data efficiency and optimization techniques, focusing on creating compact, accurate, and scalable machine learning systems. Passionate about reducing computational overhead while enhancing system performance for real-world applications.*

EDUCATION

- **Purdue University** West Lafayette, IN
PhD in Electrical & Computer Engineering; GPA: 3.66 *May 2019 - Present*
 - **Relevant Areas:** Data Efficiency, Privacy-Preserving Machine Learning, Optimization Techniques
 - **Research Advisor:** Prof. Kaushik Roy
- **Purdue University** West Lafayette, IN
MS in Electrical & Computer Engineering; GPA: 3.72 *Aug 2017 - May 2019*
 - **Thesis:** Energy Efficient Byzantine Agreement Protocols for Cyber-Physical Resilience
- **PES Institute of Technology and Science** Bangalore, India
Bachelor of Engineering in Electronics and Communications; GPA: 9.77/10.0 *Aug 2013 - May 2017*

WORK EXPERIENCE

- **Research Intern, Integrated Systems Team** Skillman, NJ
Latent AI *May 2023 - August 2023*
 - Built data annotation tools to streamline evaluation processes for anomaly detection frameworks.
 - Designed scalable systems to handle noisy datasets efficiently and ensured robust model performance.
 - Improved energy and latency metrics of internal tools for optimized machine learning workflows.

RESEARCH PROJECTS

- **Exploring Data Efficiency in Deep Learning Systems** West Lafayette, IN
Ph.D. Dissertation, Purdue University *May 2019 - Present*
 - **TRIM: Token Relevance via Interpretable Multi-Layer Attention:**
 - * Designed a forward-only, gradient/Hessian-free data selection using interpretable multi-layer attention fingerprints.
 - * Prioritized impactful samples by matching candidate tokens via hidden-state similarity.
 - * Reduced compute and memory overheads while achieving higher test accuracy than state-of-the-art methods at equal coreset budgets.
 - * Evaluated on LLaMA and Mistral across GSM8K, CommonSenseQA, and ARC-C with consistent transfer effectiveness.
 - * *Manuscript under review.*
 - **Coresets from Trajectories: Selecting Data via Correlation of Loss Differences:**
 - * Introduced *Correlation of Loss Differences (CLD)*, a scalable, gradient/Hessian-free metric using correlation of per-sample loss differences, to select influential training points.
 - * Achieved state-of-the-art on CIFAR-100 and ImageNet-1k across multiple subset budgets and architectures (ResNet, VGG, Swin Transformer), with < 1% drop under cross-architecture transfer.
 - * Cut compute and memory relative to gradient-based selection while preserving accuracy at fixed subset sizes.
 - * Established a convergence guarantee: training on CLD-selected coresets tracks full-data optimization, with deviation bounded by validation-trajectory alignment and validation approximation error.
 - * Published at **TMLR 2025**.

- **DOTIE: Energy-Efficient Object Detection Using Event Cameras**
 - * Proposed a novel lightweight detection framework leveraging event-driven camera data to minimize energy consumption.
 - * Surpassed conventional detection techniques in accuracy, efficiency, and latency metrics.
 - * Demonstrated at **ICRA 2023** and **CVPR 2023 Workshops**.
- **TOFU: Federated Learning with Data and Communication Efficiency**
 - * Developed an innovative federated learning framework to reduce communication overhead by $10\times$.
 - * Ensured privacy-preserving data sharing via encoded weight updates, validated for robustness against inversion attacks.
 - * Published at **IEEE Access 2024**.

SELECTED PUBLICATIONS

- **Nagaraj, Manish**, Sakshi Choudhary, Utkarsh Saxena, Deepak Ravikumar, and Kaushik Roy. Trim: Token-wise attention-derived saliency for data-efficient instruction tuning, 2025
- **Nagaraj, Manish**, Deepak Ravikumar, and Kaushik Roy. Coresets from trajectories: Selecting data via correlation of loss differences. *Transactions on Machine Learning Research (TMLR)*, 2025
- **Manish Nagaraj**, Chamika Mihiranga Liyanagedera, and Kaushik Roy. DOTIE - detecting objects through temporal isolation of events using a spiking architecture. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4858–4864, 2023
- Arjun Roy, **Manish Nagaraj**, Chamika Mihiranga Liyanagedera, and Kaushik Roy. Live demonstration: Real-time event-based speed detection using spiking neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4080–4081, 2023
- **Nagaraj, Manish**, Isha Garg, and Kaushik Roy. Tofu: Towards obfuscated federated updates by encoding weight updates into gradients from proxy data. *IEEE Access*, 2024

RELEVANT COURSEWORK

- Artificial Intelligence • Statistical Machine Learning • Random Processes and Probability • Linear Algebra
- Computational Models and Algorithms(DSA) • Distributed Computer Systems • Computer Networks

SKILLS

- **Programming Languages and OS:** Python, Ubuntu
- **Software Development Tools:** Docker, GitHub
- **DL Frameworks and Libraries:** PyTorch, HuggingFace, OpenCV, numpy, scipy

REFERENCES

PhD Advisor

Prof. Kaushik Roy
Purdue University,
West Lafayette, USA
kaushik@purdue.edu