RAKSHITH SHARMA SRINIVASA

Research Scientist, Meta SuperIntelligence Labs
Menlo Park, CA

✉ rakshith.sharma.s [at] gmail.com
🌐 `rakshithsrinivasa.github.io`

SUMMARY

Research Scientist specializing in LLM post-training, preference alignment, and multimodal reasoning. Four years of AI research experience with hands-on expertise in post-training algorithms, reward modeling, and RLHF pipelines. Published in top-tier machine learning venues such as NeurIPS and ICML, with multiple patents in machine learning and signal processing.

EDUCATION

Ph.D in Electrical and Computer Engineering Aug 2015 - December 2020
Georgia Institute of Technology, Atlanta, GA GPA:3.93/4.0
Advisor: Dr. Justin Romberg
Outstanding Research Award, 2020 (Center for Signal and Image Processing, Georgia Tech)
ITA Graduation Day Award, 2020

M.S in Electrical and Computer Engineering Aug 2014 - December 2020
Georgia Institute of Technology, Atlanta, GA GPA:4.0/4.0

B.Tech in Electronics and Communication Engineering July 2010 - May 2014
National Institute of Technology Karnataka, Surathkal, India GPA:9.36/10.0

EXPERIENCE

Research Scientist, Meta SuperIntelligence Labs Sept 2025 – Present
Menlo Park, CA

- **Post-training and safety research:** Developing methods for LLM post-training and evaluations for safety.

Research Scientist, Scale AI Jan 2025 – Sept 2025
San Francisco, CA

- **Multimodal AI research:** Led LLM post-training for multimodal understanding and reasoning; developed training recipes that improved performance on MMMU, MMMU-Pro, and MathVista.
- **Post-training for tool use:** Designed datasets for SFT and RL-based fine-tuning to enhance tool usage capabilities.
- **RLHF data quality assessment:** Trained reward models and applied PPO to evaluate and improve the quality of post-training datasets.
- **Benchmark design:** Developed two evaluation benchmarks to measure frontier capabilities of LLMs.
- **Post-training algorithm research:** Investigating advanced post-training methods, including rubric-based approaches, multimodal RLVR, and curriculum learning.

Senior Research Scientist, Samsung Research America (SRA) Dec 2021 – Jan 2025
Mountain View, CA

- **LLM inference efficiency:** Improved inference latency of small language models (SLMs) by **36%** using speculative decoding and multi-token prediction.
- **Vision-language pre-training:** Developed a novel contrastive training framework achieving state-of-the-art zero-shot accuracy
- **Multilingual LLM training:** Fine-tuned LLMs for **language translation** using SFT

Senior Research Scientist - IQVIA Jan 2021 – Nov 2021
Cambridge, MA

- **AI for clinical trials:** Developed ML solutions for clinical trial operations, health condition prediction and rare disease prediction
- **Fair ranking:** Designed ranking models for medical providers in clinical trials, focusing on fairness in patient selection.

SELECTED PUBLICATIONS

- C. Lee, C. Yang, J. Cho, Y.M. Saidutta, **R.S**, Y. Shen, H. Jin, 'RestoreGrad: Signal Restoration Using Conditional Denoising Diffusion Models with Jointly Learned Prior', **ICML**, Vancouver, B.C, July 2025

- Y.M. Saidutta, **R.S**, J. Cho, C. Lee, C. Yang, Y. Shen, H. Jin, 'CIFD: Controlled Information Flow to Enhance Knowledge Distillation', **NeurIPS**, Vancouver, B.C, December 2024

- **R.S**, J. Cho, C. Yang, Y.M. Saidutta, C. Lee, Y. Shen, H. Jin, 'CWCL: Cross-Modal Transfer with Continuously Weighted Contrastive Loss', **NeurIPS**, New Orleans, Louisiana, December 2023

- **R.S***, Y.M. Saidutta*, C. Lee, C. Yang, Y. Shen, H. Jin, 'To wake-up or not to wake-up: reducing keyword false alarm by successive refinement', **International Conference on Acoustics, Speech and Signal Processing (ICASSP)**, Rhodes Island, Greece, June 2023 (* - equal contribution)
- **R.S**, S. Kim, K. Lee, 'Recovering sketched low-rank matrices with a shared factor by convex programming', **IEEE Journal on Special Areas in Information Theory (Special Issue: Sensing: Fundamental Limits and Modern Applications)**, 2023
- J. Gao, **R.S**, C. Qian, L. Glass, J. Spaeder, J. Romberg, J. Sun, C. Xiao, 'STAN: Spatio-Temporal Attention Network for Pandemic Prediction Using Real World Evidence', **Journal of the American Medical Informatics Association** (JAMIA), November 2020
- **R.S**, C. Xiao, L. Glass, J. Romberg, J. Sun, 'FastGAT: Fast Graph Attention Networks Using Effective Resistance Based Graph Sparsification', **Preprint**, https://arxiv.org/abs/2006.08796
- **R.S**, M. Davenport, J. Romberg, 'Sample complexity bounds for localized sketching' **AISTATS**, August 2020 ,
- **R.S**, K. Lee, M. Junge, J. Romberg, 'Decentralized sketching of low rank matrices' **Neural Information processing systems (NeurIPS)** , Vancouver, Canada, December 2019

TALKS

- "Localized Sketching for matrix multiplication and regression", LightOn (Paris) summer seminar series, June 2020
- "Subspace learning and embedding with localized sketching" - Graduation day presentation, Workshop on Information theory and applications (ITA), San Diego, february 2020
- "Localized matrix sketching with applications to active array imaging", Spectrum Lab, Indian Institute of Science, Bangalore, India, Ferurary 2019

TECHNICAL SKILLS

- Python, Pytorch, C++, MATLAB, SQL
- Linux, macOS, Git, LaTeX

SERVICE, TEACHING EXPERIENCE

- Area Chair, AISTATS 2023, 2024
- Reviewer, NeurIPS, ICLR, ICML
- Reviewer, Transactions on Signal Processing , Transactions on Information Theory
- Session Chair, Allerton Conference, 2018
- Teaching Assistant, **Math foundations of Machine learning, Statistical machine learning**