Soumyajit Karmakar

soumyajit.karmakar.cse@gmail.com | GitHub | LinkedIn | Google Scholar
soumyajitkarmakar.github.io

Education

University of California San Diego 2024 - 2026
M.S. in CS | **GPA: 3.92/4.0**
Advisor: Mikhail Belkin
Thesis: Learning Relations Between Steering Vectors

Indian Institute of Information Technology, Guwahati 2019 - 2023
B.Tech. in CSE | **CGPA: 9.35/10** | **Silver medalist**
Co-advisor: C. V. Jawahar, IIIT Hyderabad

Experience

Belkin Lab, UC San Diego Sep 2024 - Present
Research Assistant
Advisor: Mikhail Belkin

- Developed linear transformations for **LLM activation steering** across semantic opposites, layers, and translation tasks, generalizing to held-out concepts.
- Showed on Llama 3.1 8B, Gemma 2 9B, and Mistral NeMo 12B that a single learned linear map per model converts a steering vector into one for its antonym (e.g., cheerful → gloomy) on held-out pairs without any data for the target concept, **matching directly extracted vectors**.

Fujitsu Research, Bengaluru Sep 2023 - Aug 2024
ML Engineer
Collaborator: R. Venkatesh Babu, IISc Bengaluru

- Led development and deployment of **Prompt-ReID**, which re-identifies people and vehicles across cameras by matching images to one learned CLIP text embedding per identity instead of individual photos, beating the existing SotA (CLIP-ReID) by **2.7 points** Rank-1 on MSMT17 (**91.1%**) and out of distribution by an average of **10.1 points**.
- Built the distributed PyTorch training codebase for Prompt-ReID, which ran on 3 clusters with different schedulers and GPU generations and scaled to **32 GPUs** across **4 nodes**, cutting a full training run from 18 hours on a single node to 5 hours.

University of North Carolina at Charlotte Aug 2022 - Mar 2023
Research Intern
Advisors: Srijan Das and Michael S. Ryoo

- Co-designed **Self-Supervised Auxiliary Task (SSAT)** training, jointly optimizing classification and masked autoencoder reconstruction with a shared ViT encoder and **no added inference overhead**.
- Benchmarked ViT, CvT, and Swin on **10 datasets**; raised ViT-S CIFAR-100 top-1 accuracy from **54.08% to 73.37% (+19.29 percentage points)** under equal 100-epoch training.

CSIR-CEERI, Pilani May 2022 - Jul 2022
Research Intern at the Advanced Information Technologies Group (AITG)
Advisor: Sanjay Singh

- Built a few-shot classification framework that fuses features from frozen pre-trained CNNs (ResNet-50, EfficientNet-B5, DenseNet-201) through a lightweight convolutional ensembling block.
- Achieved **92.30%** 5-way 5-shot accuracy on a drone-captured power-line defect dataset, and outperformed the library-learner baselines on 5 of 6 standard benchmarks.

Achievements

Best student submission award, Parameterized Algorithms and Computational Experiments (PACE) 2022 Heuristic Track (Directed Feedback Vertex Set), with solver FEDRER.

Publications

- **Soumyajit Karmakar**, N. Mallinar, M. Belkin. Linear Transformations Between Steering Vectors in Large Language Models. (under review at AAAI 2027)
- **Soumyajit Karmakar**, S. Ramasubramanian, H. Rangwani, Y. Aoki, G. Suzuki, R.V. Babu. Prompt-ReID: Prompt Based Inference for Robust Re-Identification. Fujitsu Research, 2024 (Technical Report).
- S. Das, T. Jain, D. Reilly, P. Balaji, **Soumyajit Karmakar**, S. Marjit, X. Li, A. Das, M.S. Ryoo. Limited Data, Unlimited Potential: A Study on ViTs Augmented by Masked Autoencoders. **WACV 2024**.
- **Soumyajit Karmakar**, A. Banerjee, P.S. Gidde, S. Saurav, S. Singh. Convolutional Ensembling based Few-Shot Defect Detection Technique. **ICVGIP 2022**.
- A. Jain, S. Agarwal, N. Agrawal, **Soumyajit Karmakar**, S. Swain. Feedback vertex set using Edge Density and REmove Redundant (FEDRER): A heuristic solver for finding a feedback vertex set in a directed graph. **PACE 2022 solver description**.

Projects

Model Insurance
Undoing Edits to Trained Models with Little Protected Information
Independent research project
Report

- Formulated the cost of guaranteeing that any edit from a given class can be undone, and showed it is set by the chromatic number of a graph of confusable models, which error-correcting codes achieve up to constant factors.
- Proved that a record of the model's outputs lowers this cost. It drops to zero for sparse bias edits in one-hidden-layer ReLU networks once the record covers every neuron's isolating input region. For rounded final-layer outputs, it drops by about **2 bits** each time the record doubles, with matching lower bounds.

Meaningless Tags, Meaningful Impact
Formatting Decides Context-Memory Conflicts in LLMs
Independent research project
Report

- Found that wrapping a false, source-attributed claim in <document> tags flips Gemma 4 12B from 0% to 100% adoption of the false answer. That matches an explicit instruction to adopt it, and even a meaningless tag (<qzx_block>) reaches **75–79%**.
- Designed 16 controlled wrappers over fixed content to separate delimiter syntax from tag semantics. The wrapper rankings replicate on Qwen 3.6 27B (Spearman $\rho = 0.79$–0.88).

Service

Reviewer: AAAI 2027, COLM 2026, ICVGIP 2022.
Organizing: COLT 2026 organizing team (program logistics); TILOS workshops at UCSD.
Mentoring: Mentored five incoming graduate students through UCSD's Grad Pals Mentorship Program.
Open source: Contributor to CompilerGym.

Technical Skills

Languages & Tools: Python • C++ • PyTorch • JAX • Git • LaTeX.
ML & AI: LLM Interpretability • Activation Steering • Vision Transformers • CLIP • Masked Autoencoders • Self-Supervised Learning.
Methods: Linear Probing • Recursive Feature Machines • PCA / SVD • Low-Rank Approximation • Hessian Spectral Analysis.