Sathyaprakash Narayanan

snaray17@ucsc.edu
https://satabios.github.io/

Education

- **University of California, Santa Cruz** — Sep 2022 - June 2024
  Masters in Electrical and Computer Engineering

- **Anna University; RMD Engineering College, Chennai** — June 2014 - April 2018
  Bachelors of Engineering, Electronics, and Communication

Work Experience

- **AI Engineer** — March 2025 - Present
  Qualcomm Inc.
  San Dieogo, California, USA

  - End-End Development of Compute-Friendly ADAS based BEV models
  - Research and Development experience in the following areas:
    * Optimization of Deep Learning Kernel, Compiler Tools, Model Efficiency Tools
    * Pruning, Quantization, NAS, Knowledge distillation
    * Optimizing and deploying inference on Qualcomm processors

- **Machine Learning Scientist II** — Oct 2021 - April 2022
  Lytx India Tech Inc.
  Bangalore, Karnataka, India

  - End-End Development of Compute-Friendly ADAS based ML/DL models
  - Research and Development experience in the following areas:
    * Model compression and NAS techniques
    * Knowledge distillation, Pruning, Quantization
    * Optimizing and deploying inference on various embedded processors
    * Signal Processing on mmmwave Radars

- **Research Associate** — Dec 2017 - Sep 2021
  NeuRonICS Lab, DESE Department,
  Indian Institute of Science, Bangalore

  - Real-Time Object Detection and Localization in Compressive Sensing Video — **ICIP 2021**
  - N-HAR: neuromorphic event-based human activity recognition system — **ISCAS 2019**
  - Proto-object based visual saliency model on NVIDIA Jetson TX — **ISCAS 2019**
  - n-EAR: Neuromorphic Ego Motion - Vehicle Activity Recognition — **US Patent App. 17/377761**
  - System and Method for exhale controlled AAC device for IOT — **IN Patent App. 201641044496**

Open Source Projects

- **sconce: Model Compression and Deployment Package** — **Author**

  - Built a one-stop compression package, that supports:
    * Pruning, Quantization and NAS
    * Deployment in ONNX, C/CUDA code generation
    * Sparsity Engine and Support for Inference level Optimizations

- **snntorch: Gradient Based Spiking Neural Networks** — **Contributor**

  - Built Pruners for Spiking Neural Networks
  - Built support for model compression for snn models on sconce

Technical Skills and Interests

LLM/Fine-Tuning, Machine Learning, Deep Learning, Computer Vision, Model Compression, Signal Processing, Pruning, Quantization, MLOps/AIOps, Git, MATLAB, Python, C/C++, CUDA, Triton

Positions of Responsibility

- **Reviewer**
  - MLSys 2026
  - WACV 2019, 2020, 2022
  - ICIP 2019, 2020, 2022, 2023, 2025
  - TPAMI IEEE Transactions on Pattern Analysis and Machine Intelligence