

Sumanth Gurram

San Francisco, CA | sumanthgurram.com | linkedin.com/in/sumanth-gurram | github.com/sumanthgenz

Education

University of California, Berkeley BS in EECS, BS in Business Aug 2019 – May 2023

- Technical GPA: 3.9/4.0
- M.E.T. Dual Degree Program
- IEEE Eta Kappa Nu (top 25% of EECS)

Experience

Member of Technical Staff, Decagon – San Francisco, CA March 2025 – Present

- Building agentic AI for enterprise customer support

Software Engineer, Zip – San Francisco, CA June 2023 – Feb 2025

- Drove AI features (Assistant, similar vendors) and enhanced core product (team management, user transitions)
- Reduced search overhead by 12X on core dashboards; improved p50 and p75 latency for core mutations by 40%

Research Engineer Intern, Meta AI – Menlo Park, CA July 2022 – Sept 2022

- Built support for vision (gaze-tracking) and speech as multi-modal inputs for interactive image segmentation
- This was research for Segment Anything at FAIR; 3rd most cited AI paper worldwide in 2023

Software Engineer Intern, Nuro – Mountain View, CA April 2022 – July 2022

- Built cost tracking framework, led data migration to GCP and created synthetic traffic testing jobs
- This helped scale NuFS, Nuro's distributed file system, to store over >100 PB self-driving data

Machine Learning Intern, Apple – Cupertino, CA May 2021 – Aug 2021

- Created and trained 3D object pose-estimation model with simulation data for robot vision
- Presented to VP of Hardware to help automate device testing/QA for Apple Watch

Machine Learning Researcher, BAIR, SkyLab – Berkeley, CA June 2020 – June 2023

- Advised by Prof. John Canny for multi-modal self-supervised video understanding using transformers
- Advised by Prof. Ion Stoica (co-founder of Databricks) for multi-cloud ML training infrastructure

Publications

LAVA: Language Audio Vision Alignment for Contrastive Video Pre-Training May 2022

S. Gurram, A. Fang, D. Chan, J. Canny
2022 ICML Workshop

Projects

Arrow AI

- Developed a platform to evaluate prompts and unit test LLMs using natural language test suites
- Built for startup I co-founded in 2022 Soma Capital Fellowship; turned down further funding to resume college

Technologies

Languages: Python, Java, C++, C, Go, Typescript, Javascript, HTML/CSS, Solidity, RISC-V, Intel x86

Tools: PyTorch, TensorFlow, OpenAI API, Elasticsearch, GCP, AWS, Kubernetes, PostgreSQL, GraphQL, React.js, DataDog, Kibana, Unity, LaTeX