Sarthak Sethi

San Jose, CA — 925-895-0285 — sarthakluv@gmail.com — linkedin.com/sarsethi — gitHub.com/1300Sarthak — Portfolio Site

Education

San Jose State University **Expected Graduation: Dec 2026**
Bachelor's of Science in Computer Science and Applied Math
Awards: *Dean's Scholar (2x) & 4x Hackathon Winner* — ***Course Work:*** *DSA, Databases, Machine Learning, OOP*

Skills

Languages	Python, Java, Typescript, JavaScript, C#	**GenAI**	LLMs, Agentic AI, Fine-tuning (LoRA)
Frontend	React, Next.js, JavaScript, HTML/CSS	**GenAI Infra**	RAG, MCP, Vector databases
Backend	Springboot, Django, Flask, .NET, NGINX	**Cloud**	AWS (Lambda, EC2, S3), Kubernetes, Docker
Databases	SQL(PostgreSQL, Oracle), NoSQL(MongoDB)	**DevOps/Tools**	CI/CD, Git, Postman, Jira, Agile

Experience

Incoming - Software Engineering Intern **Sep 2025 – Dec 2025**
Lighthouse Observatory (Y-Combinator 2025) *Remote*

Software Engineering Intern **May 2025 – Aug 2025**
Gilead Sciences *Foster City, CA*

– Led end-to-end build of a Model Context Protocol (MCP), AI agent bot on an in-house fine-tuned LLM with secure intranet integration that surfaces real-time Identity and Access Management (IAM) status via chat and daily alerts, cutting manual access tickets 37%.

– Built and deployed a Python security automation suite using Selenium, Django, and unittest to test file uploads and enforce CI gates, blocking 15+ high-risk defects before release.

– Shipped an analytics platform by migrating 20,000+ legacy records into Oracle with Python ETL, exposing a secure REST API, and delivering Power BI dashboards with row-level security and scheduled refresh, cutting reporting cycle time 60%.

Data Engineering Intern **Jan 2025 – Mar 2025**
Cequence Security *Santa Clara, CA*

– Built a Python and SQL REST-traffic analytics pipeline with 1h, 24h, etc. windowed aggregations and alerting to track per-endpoint volume, errors, and latency, improving system efficiency 15%.

– Integrated Open Policy Agent (OPA) with Nginx using Rego policy-as-code to enforce API access control and request validation, keeping P99 gateway overhead under 5 ms.

– Built a Python solution integrated into CI/CD to analyze REST API request behavior and auto-flag anomalies, reducing mean time to detection by 45%.

Machine Learning Research Assistant **Jan 2025 – Present**
MICoSys Lab @SJSU *San Jose, CA*

– Researching adversarial machine learning on LLMs, under the guidance of Dr. Sengupta, designing and executing prompt-based jailbreak evaluations pre- and post-fine-tuning to assess model robustness and alignment.

– Fine-tuned small-scale LLaMA models using LoRA (Low-Rank Adaptation) adapters on novel datasets, improving performance and analyzing generalization across unseen tasks.

– Explored unlearning techniques by programmatically removing and modifying LoRA adapters; decomposed delta weights to identify and interpret shifts in model behavior and learned representations.

Projects

Co-Founder - Machine Learning Based Smart Data Analyst and Visualization SaaS **Github Link**

– Built and own full-stack AI analytics SaaS (LangChain, React, Supabase, AWS) for CSV queries, SQL generation, and instant visual/statistical insights
– Architected scalable cloud deployment supporting high concurrency, secure data handling, and sub-second queries

RAG/MCP Based Government Chat-bot — 1st Place Winner at Microsoft Azure Hackathon **Link**

– Location-aware Model Context Protocol with integrated government form autofill and Retrieval-Augmented Generation (RAG) pipeline