+1 (412) 287-2364 sayyampe@andrew.cmu.edu

SURYA GANESH

linkedin.com/in/suriya-ganesh
suriya.cc
Google Scholar

EMPLOYMENT

Senior Systems Engineer NimbleEdge Jan 2023 - July 2023

On-Device ML

- Implemented Federated Learning (FL) Research such as FedAvg, Model Fine-tuning in the edge, Client Selection Polling, Secure Aggregation, etc.papers into Deployable and maintainable applications.
- Designed platforms with 100M MAU. Systems I built were handling 20K requests per second at peak load. We did 50M inferences a day on EdgeML deployments.
- Led Architecture decisions on embedded edgeML SDK built on top of Onnx runtime.
- Mentored Junior engineers and interns on deliverables and Goals.

Software Development Engineer II

Skyflow

Oct 2021 - Dec 2022

Governance Engine Team

- Optimized the governance engine reducing p99 response time from 150 ms to 5ms at 15K RPS by leveraging hierarchical caches and cache invalidation algorithms.
- Redesigned control and data planes, from a single tenant architecture to a multi-tenant architecture. With data residency across multiple regions. Reducing Server cost by 67% annually.
- Migrated from an Arangodb based backend to Mongodb, with 80 Million records coming to about 1TB data, with zero downtime To improve reliability and performance.

Platforms Engineer Acko Insurance July 2019 - Oct 2021

Platforms Team

- Managed more than 300 Releases over a period of 7 months, while maintaining zero down time.
- Collaborated with 9 different teams and migrated their services into a common platform. Improving SLAs
 across different Lines of Businesses

EDUCATION

MS - Privacy Engineering Carnegie M

Carnegie Mellon University

Aug 2023 - Dec 2024

GPA 3.78/4, Coursework: Deep Learning, Prompt Engineering, AI Venture Studio, etc.

MTech - Software Engineering

BITS Pilani

Dec 2020 - Dec 2022

GPA 3.9/4, Coursework: Distributed Systems, Database Design etc

BE - Instrumentation Engineering

Anna University

April 2015 - April 2019

GPA 3.9/4, Coursework: Control System, Digital Signal Processing, VLSI Design, etc.

PROJECTS

Core Contributor Speaksharp.io Mar 2024 - Present

- Built speaksharp.io, a digital language and communication coach converting your personal daily conversations to learning opportunities, using Python, voice diarization, prompt engineering, etc.
- Top 10% for YC Summer '24 batch.

SKILLS

- Languages: Rust, Golang, C++, Java, python, javascript
- Tools/Frameworks: Onnx, Akafka, K8s, Docker, EKS, Azure, gRPC, AWS, react-native, Ray, Kotlin, Elasticsearch

ACHIEVEMENTS AND PUBLICATIONS

- SPARK grant for Entrepreneurship, CMU '24
- Privacy Engineering Scholarship, CMU '23
- publication: "Through the Lens of LLMs, Unveiling Differential Privacy Challenges", Usenix PEPR '24