

Sai Ravi Teja Gangavarapu

(412) 251-7161 | sairavig@cs.cmu.edu | Pittsburgh, PA | floaredor.vercel.app

EDUCATION

Carnegie Mellon University

Master of Software Engineering - Scalable Systems (Statistics, QA, Design Patterns, Prod Mgmt)

Pittsburgh, PA

University of Florida

Senior Cert + Master's coursework in Computer Science (Algorithms, UX, Graphics - 4.0)

Aug 2025 – Dec 2026

Gainesville, FL

Jan 2024 – May 2024

SKILLS

Python, C++, Go, TypeScript, SQL, CUDA, PyTorch, TensorFlow, React, Next.js, FastAPI, Langgraph, Docker, Kubernetes, AWS, PostgreSQL, Redis, ArgoCD, OpenAI, Cursor (Power User), Test and Spec-driven Development, Brownfield Refactoring, QA & Testing, MathAcademy Mathematics for ML, Deep Learning, Machine Learning, Statistics, Linear Algebra, MPI, HPC

EXPERIENCE

Founding Software Engineer — Tapsta

Jul 2024 – Jun 2025

- Architected end-to-end full-stack social rewards mobile application using React, React Native, FastAPI, and PostgreSQL.
- Designed database architecture with 25+ schemas and optimized client-side operations with caching and debouncing, achieving 30% faster response times. Worked with product team to redesign onboarding flow, driving 40% user retention.
- Orchestrated Plaid API integration alongside REST API creation, establishing secure bank connectivity and enabling automated cashback processing, plus ACH transfers for over 1,500 application users.
- Made containerized CI/CD pipeline with Docker, AWS, and ArgoCD that reduced deployment time by 70%.
- Led development team of 3 SDE interns, establishing code review processes that contributed to **30%** faster feature delivery.
- Built a semantic people search engine for alumni networks using embedding-based similarity using LangGraph.

Research Assistant - Fan Lab, University of Florida

Feb 2024 – Jun 2024

- Applied genomic models for rare disease prediction using **transformer architectures**, processing **15,000+ DNA sequences**. Developed multi-layer perceptron models using protein embeddings (**ESM3**) analyzing RNA-binding proteins.

SDE Intern — Hashira.io

Apr 2023 – Dec 2023

- Taught LLMs to use APIs (incl. function calling), fine-tuning LLaMA/BERT in PyTorch to build a natural-language interface for payroll, achieving 90% natural language to API accuracy and reducing API integration time by 25%.
- Implemented full-stack crypto analytics platform using React, FastAPI, Postgres, Pandas, and Golang with real-time monitoring dashboard and a microservice for a leaderboard system for garden.finance, boosting user engagement by 30%.
- Improved backend performance using Go and PostgreSQL optimizations, supporting \$150M+ trading volume over 30 days with 20% faster queries on garden.finance.

Research Assistant - Mahindra Ecole Centrale

Dec 2022 – Jan 2025

- Published research on emotion-targeted music generation using FFT, differential evolution, F C-means, self-organizing maps achieving 85% accuracy in classification (IEEE CEC 24). Implemented ALI-GAN model with t-SNE and PCA for clustering analysis on 1000+ samples, achieving unsupervised music genre classification. Built MIR pipelines extracting 100+ features.

PROJECTS

Project RECON | Raspberry Pi, OpenMPI, GlusterFS | \$2000 funding

- Architected 8-node Raspberry Pi 4B **compute cluster** with distributed storage, LDAP authentication, and Slurm job scheduling, serving 400+ students for coursework.

OneAIClick.com | Python, React.js, HuggingFace, PEFT/LoRA (*Built no-code/low code tool*)

- Developed an LLM fine-tuning abstraction tool for HuggingFace models that cut development cycles by 60% and reduced boilerplate code by 75%, enabling rapid idea validation and full data privacy.

CUDA Ray Tracer & Audio-Reactive Visualizer and Graphics Engine | C++, CUDA, OpenGL

- Built GPU-accelerated ray tracing system achieving **1,600x speedup** (7.5s to 0.0045s per frame) over CPU implementation, with real-time FFT-based audio reactive 256-agent boids simulation. Also, built modular 3D renderer in C++ and OpenGL

Interpretable & Controllable Text-to-Audio Synthesis | PyTorch, CLAP, TorchSynth, Sound Design

- Built a text-conditioned synthesizer controller maximizing text–audio embedding similarity. Also, did inverse synthesis.

RootCause Auto-Remediation Agent | Next.js, OpenTelemetry, MongoDB, FastAPI, Python, Gemini

- Engineered an **OpenTelemetry** pipeline to parse traces and deduplicate alerts in MongoDB, reducing noise by **45%**.
- Built a self-healing agent that auto-generates **syntax-validated (AST)** fixes via Gemini, slashing MTTR by **75%**.

LEADERSHIP AND ACHIEVEMENTS

President, Enigma, the Computer Science Club (2021-2023): Conducted technical workshops on Gamedev, ML, Linux, heading outreach initiatives increasing the club membership by 40%, reaching over 2000 students. Collabs - Ubisoft, NVIDIA.

1st Place WaffleHacks 2024 (320+ participants): Built FlashFocus - a Chrome extension to block distractions and turn them into learning with AI flashcard quizzes using RAG system. (React, FastAPI, SQL, Groq)

1st Place Talentmapp23 (300+ participants): Made an intelligent task-tracking app with LLM suggestions - React, GPT, Flask

1st Place Noderunner Hackathon 2023 (50 teams): Engineered distributed cluster implementing Raft consensus protocol from scratch (Python, Flask, Multithreading)

Buildspace S5 30+ users - Desktop app, AI sample packs from songs with stem separation, drum extraction & flips for musicians

Publications - "Emotion Aligned Music Composition from Sound Fundamentals Using Differential Evolution" - IEEE CEC 24