

Sai Ravi Teja Gangavarapu

+1 (412) 2715161 | sairavig@andrew.cmu.edu | Pittsburgh, PA | floaredor.vercel.app

EDUCATION

Carnegie Mellon University

Master's in Software Engineering - Scalable Systems

Aug 2025 – Dec 2026

Pittsburgh, PA

University of Florida

Senior Certificate + Master's coursework in Computer Science, CISE Department

Jan 2024 – May 2024

GPA: 4.0/4.0

- Courses: Analysis of Algorithms, Advanced Data Structures, Computer Graphics, UX Design

Mahindra École Centrale

B.Tech, Computer Science and Engineering

2020 - 2024

Hyderabad, India

RESEARCH EXPERIENCE

Fan Lab, University of Florida

Research Assistant — Language Models, Supervised ML, Neural Networks

Feb 2024 – Jun 2024

Gainesville, FL

- Conducted research under Dr. Xiao Fan on genomic foundational models using nucleotide variance techniques for rare disease prediction
- Developing neural network system for RNA-Protein interaction prediction to aid in cancer diagnosis.

Michigan State University

Research Assistant — Evolutionary Algorithms, signal processing, PyTorch, MIR, music

Sept 2024 – Present

Remote

- Working with Dr. Kalyanmoy Dev, Prof. Arya Kumar Bhattacharya on optimization for emotional music transitions
- Developing therapeutic music generation systems using the iso principle for mood transition
- Investigating **multi-objective optimization** for raga and rasa(emotion)-based music generation

Mahindra École Centrale

Research Assistant — PyTorch, signal processing, MIR, music, deep learning, feature engineering

Oct 2022 – Sept 2024

Hyderabad, India

- Published in **IEEE WCCI 2024 (CEC)**: Developed emotion-targeted Hindustani music generation system
- Designed and implemented pipelines for extracting and filtering 100+ time- and frequency-based features using Music Information Retrieval (**MIR**) techniques.
- Implemented **Differential Evolution** to optimize Fourier transform coefficients with fitness functions based on **self-organizing maps** (SOM) and **fuzzy c-means** clustering, achieving 90% classification accuracy.
- Previously, researched **unsupervised genre classification** by building a **Generative Adversarial Network** using PyTorch to encode Indian Classical music spectrograms and perform **t-SNE** clustering analysis.

IIT Palakkad - ACM Program

Research Trainee — Docker, SLURM, OpenMP

Jun 2023 – Jul 2023

Palakkad, India

- Implemented parallel scientific simulations using **MPI** and **OpenMP** on ParamVidya supercomputer
- Implemented distributed computing applications using **Docker** and **Kubernetes**

WORK EXPERIENCE

Tapsta

Lead Software/Founding Engineer — React Native, JS, Python, FastAPI, PostgreSQL, Docker, Git

Aug 2024 – May 2025

Remote

- Led development team of 3 building social media app connecting students with local merchants (**1000+ users**)
- Designed architecture from the groundup with 20+ schemas, shipped **55,000+ lines of code** across stack including mobile app, website, backend and banking integrations
- Implemented CI/CD pipelines, Docker containers, and Optimized database queries and client-side operations (e.g., caching, debouncing) for efficiency and scalability.
- Mentored 3 SDE interns, accelerating project velocity and code quality
- scaled up the servers to handle 10000+ users in anticipation
- Additionally, built an AI-first people search engine for college alumni centres.

OneAiclick.com

Co-Founder/Founding Engineer — LORA, PyTorch, FastAPI, Nextjs, React, Docker, AWS, sockets

July 2024 – Aug 2024

Remote

- Building private LLM pipeline abstraction tooling for easy fine-tuning and deployment of language models
- Developed a no-code GUI for fine-tuning open-source HuggingFace LLMs, enabling rapid and secure model deployment while maintaining full data privacy, providing a cost-effective alternative to hiring AI specialists for testing ideas.
- Developing system for secure model deployment and data privacy for agentic RAG and low parameter fine-tuning workflows. Speeding up validation of fine-tuning on custom datasets by 2X over traditional methods.
- Gained entrepreneurial experience through customer development and product iteration

Harvested Robotics

Apr 2024 – May 2024

Software Developer — React, Next.js, TailwindCSS

Remote

- Developed responsive landing page for agricultural robotics startup focusing on automated weed control solutions

Society of Software Developers, UF

Feb 2024 – May 2024

Software Development Engineer — Nextjs, FastAPI, Postgresql, Langchain, ANNs

Gainesville, FL

- Developed **Co-Write**, an open-source AI-powered learning platform that provides targeted assistance to students within professor-defined boundaries, enhancing the learning experience while promoting academic integrity.
- Built the full-stack application using **Next.js**, **Tailwind CSS**, **FastAPI**, **PostgreSQL**, and **LangChain**.
- Designed and implemented key features, including a custom **Retrieval-Augmented Generation (RAG)** model, assignment creation with AI limitations, and a **text classifier** for determining assistance types.
- Achieved rapid prototyping by completing a fully functional prototype in under 24 hours.

Catalog.fi

Apr 2023 – Dec 2024

Software Development Engineer Intern — JS, PyTorch, Golang, FastAPI, Postgresql, AWS

Hyderabad, India

- Fine-tuned language models (LLaMA, BERT, etc.) and developed a custom language model that maps natural language inputs to API calls with precise parameters, enabling seamless integration of LLMs with APIs. **Taught LLMs to use APIs.**
- Conducted research and built proof-of-concept projects leveraging machine learning to enhance user experience.
- Implemented a real-time data analytics program using React JS, FastAPI and MongoDB to monitor crypto token volume, providing actionable insights. Deployed a dashboard for monitoring metrics on Garden.finance
- Designed and developed a scalable leaderboard and rewards system using Golang and PostgreSQL, boosting user engagement by 30%. Integrated atomic swaps and deployed the system on AWS EC2 instances.
- Contributed to facilitating \$150M+ in trading volume over 30 days by solving tickets and optimizing database query speeds by 20%.

TECHNICAL PROJECTS

Audio Technology Projects

2021 – Present

- Developed Samplebox (**30+** users): a desktop application that generates complete sample packs from songs, including stem separation, drum hit extraction, and AI-powered generative variations based on music theory rules. (tools: electron js, fastapi, Librosa)
- Built text-to-synthesizer parameter mapping system using PyTorch and LAION-CLAP model
- Hosted AI-generated audio-visual art showcase by web scraping landscape images training a GAN model on DGX-1.
- Implemented **LSTM**-based symbolic music generation system
- Created a basic synthesizer and an FX plugin using C++ **JUCE** framework

Graphics & High-Performance Computing

2023 – Present

- Built **CUDA**-accelerated ray tracer with FFT-based audio reactivity of Boids simulation parameters
- Built a modular 3D renderer using C++ and **OpenGL**
- Implemented various assembly-level optimizations using SIMD instructions.
- Developed a cryptocurrency/blockchain system from scratch in Python.

Full-Stack Applications

2022 – Present

- Created **Resumesmith** (**300+** users): AI-powered resume format converter
- Built **RateMUProfs**, a course review platform with exam paper repository, featuring JWT authentication, **load balancing**, and rate limiting. Implemented using React.js, FastAPI, and MongoDB; received a perfect grade in Software Engineering course.
- Implemented **RainPod**: Multi-agent AI podcast system that is interactive where users can pause and pose questions to the AI podcast hosts.

Systems & Infrastructure Projects

2023 – Present

- Contributed to Project **RECON** an 8-pi compute cluster (\$2000 funded by the university) implementing VLANs, GlusterFS distributed storage, and Slurm with **OpenMPI** serving **400+** students
- Developed **NUMA**-aware concurrent hashmaps reducing remote cache misses by 40% on dual-socket and 100% on quad-socket systems
- Implemented custom load balancer and service discovery system for distributed applications
- Developed network protocol implementations including **CRC** and **ARQ**

LEADERSHIP & ACHIEVEMENTS

Enigma Computer Science Club

Dec 2021 – Jun 2023

President

Mahindra University

- Organized Generative AI Hackathon with NVIDIA and Game Jam with Ubisoft
- Hosted an AI visual DJ set and technical workshops on Digital imaging, Linux, cryptography, API development, and game development reaching 2000+ students

ACHIEVEMENTS

- First Place, **WaffleHacks 2024** (320+ participants) Built a Chrome extension and a Retrieval-Augmented Generation (RAG) system to transform distractions into personalized learning opportunities in under 24 hours.
- First Place, **Aether Game Jam 2024** Developed a top-down horror game in Godot, featuring an A* search algorithm, custom sound design, and procedural map generation in under 2 days.
- First Place, **Noderunner Hackathon** (50+ participants) Implemented the distributed Raft consensus protocol as a solo programmer in under 18 hours.
- First Place, **Talentmapp Hack4Hire Hackathon 2023** (300+ participants) Developed a task-tracking application using MongoDB, Flask, React.js, and ChatGPT. Designed the architecture, implemented a secure backend, and deployed the app in 8 hours. Incorporated LLMs to provide intelligent task suggestions based on user history.
- Finalist, **NVIDIA Student Ambassador Program**.

TECHNICAL SKILLS

Languages: Python, C++, Go, JavaScript/TypeScript, Java, Rust

Frameworks: PyTorch, FastAPI, React/React Native, Next.js, Spring Boot, gRPC

Systems & Tools: Docker, Kubernetes, AWS, Git, Linux, CUDA, OpenMP, Kafka

Areas: ML, ML-interfaces for audio arts, Distributed Systems, Cloud Computing, HPC, Full-Stack Development, System Design