

GOURIBHATLA SAI VIDITH



+91 8328258721



saividith396@gmail.com



saividith.tech



sai-vidith



SaiVidith

SUMMARY

Penultimate-year CSE student with hands-on experience across software development, cloud infrastructure, and systems integration. Interned at INCOIS (government research) building production data pipelines; independently designed and operate a **Ray-based GPU cluster** across 6–8 physical nodes; built and deployed containerised applications using **Docker, Nginx, Redis, and PostgreSQL**. Experienced in **Python, JavaScript, Linux, Git, and REST APIs** — comfortable working across development, testing, database, and infrastructure layers.

EDUCATION

VNR Vignana Jyothi Institute of Engineering and Technology
B.Tech - Computer Science and Engineering (Data Science), CGPA: 8.87

2023 – 2027
Hyderabad

EXPERIENCE

AI/ML Project Intern, INCOIS (Indian National Centre for Ocean Information Services) Mar – May 2025

- Designed and implemented an end-to-end **data pipeline** in Python to ingest, process, and structure **17 years** of domain-specific documents (2008–2025); handled ETL across **100+ large-scale PDFs** including OCR fallback for scanned documents and structured metadata storage.
- Automated document processing workflows using **Python and API integrations** to eliminate **40+ hours/month** of manual effort; produced project and process documentation, collaborated directly with domain scientists to validate outputs and gather requirements iteratively.
- Maintained and supported the pipeline end-to-end — diagnosed format-specific extraction failures, implemented fixes, and ensured consistent output quality across diverse document types spanning 17 years of records.

PROJECTS

VJ GPU Cluster — Distributed Compute Infrastructure ↗ — *Ray, Python, Linux, Docker, FastAPI*

- Architected and operate a **production compute cluster** using the Ray framework across **6–8 physical nodes**; implemented **server/system administration** including node configuration, resource allocation, job scheduling, and real-time monitoring across heterogeneous hardware.
- Configured **NFS (Network File System)** for shared storage across all cluster nodes; set up **Ray head and worker nodes** with SSH-based passwordless authentication, environment synchronisation, and fault recovery; implemented CUDA-aware resource allocation enabling **single-GPU and multi-GPU (8-GPU) parallel job execution** with live job status monitoring via the web UI.

Axon — AI-Powered Second Brain ↗ — *FastAPI, ChromaDB, Firebase, Groq, Gemini, Next.js*

- Deployed RAG knowledge system: **FastAPI** embeds docs via **all-MiniLM-L6-v2** into **ChromaDB**, routes queries across **Groq (LLaMA 3.3)** and **Gemini** with Firebase Auth, per-user Firestore profiles, and a rate-limited public API — live on Render + Vercel.

AI Business Tool Kit ↗ — *Node.js, MCP, Docker, MongoDB, Redis, Nginx, WebSockets*

- Containerised **multi-agent AI automation platform** using **MCP**: Node.js API gateway routes tasks to an Agent Registry dispatching work across **5 Docker microservices** (Website Generator, Marketing Agent, Image Generator, Analytics, VoiceBot) — enabling end-to-end automated AI workflows.
- Redis** queues for inter-agent routing, **MongoDB** persistence, **Nginx** reverse proxy with rate limiting, non-root containers with per-agent CPU/memory constraints, and **WebSockets** for real-time updates.

Emergency Vehicle Detection System ↗ — *YOLOv8, PyTorch, OpenCV, Python*

- Built a real-time object detection pipeline using **YOLOv8** and **PyTorch**, training a custom model to identify emergency vehicles in live video streams; optimised inference speed and packaged the model into a production-ready system with application-level decision logic.

TECHNICAL SKILLS

Databases:	PostgreSQL, MongoDB, MySQL, Redis, ChromaDB
Languages:	Python, JavaScript, TypeScript, SQL, Bash, Java (basics)
DevOps & Cloud:	Docker, Kubernetes (basics), Linux, Git, GitHub, Nginx, REST APIs, AWS (basics), GCP
Web & Backend:	FastAPI, Node.js, Express.js, REST APIs
Analytics & Misc:	NumPy, Pandas, PyTorch, HuggingFace, Splunk (basics), Data Visualisation
Systems & Tools:	Git, Docker, Nginx, Linux, WebSockets, MCP, JIRA

LEADERSHIP & ACHIEVEMENTS

- Tech Lead, VJ DataQuesters (Website)** — Led technical initiatives for a **200+ member** community (50+ technical); shipped production features to the club platform and mentored junior members on engineering best practices.
- Runner-up, Technovista Hackathon** — Delivered a working AI-based system under time constraints; demonstrated end-to-end problem solving and team execution in a competitive setting.
- Ranked 47th / Top 100** in competitive programming contest; solved **160+ problems** across arrays, graphs, dynamic programming, and algorithms on LeetCode.

CERTIFICATIONS

AI Fundamentals — IBM SkillsBuild/Cisco — Prompt Design in Vertex AI — Google Cloud — SQL — HackerRank