

Sanjay Samala

☎ +91 9704665956 ✉ sanjaysamala4100@gmail.com 💻 sanjaysamala 🌐 sanjaysam410 📁 sanjaysamala 🌐 Sanjay-Portfolio

OBJECTIVE

Python developer with experience in building scalable backend applications, data processing pipelines, and full-stack web solutions. Skilled in developing real-world systems using Python, REST APIs, databases, and modern frameworks, with additional experience in Generative AI and Retrieval-Augmented Generation (RAG) applications.

EDUCATION

Matrusri Engineering College

Bachelor of Engineering in Computer Science and Engineering

Nov. 2022 – June 2026

CGPA: 7.04

EXPERIENCE

AI Developer Intern — Viswam.ai

Oct 2025 – Feb 2026

Low-Resource Language AI Initiative (Meta x Viswam.ai x IIIT Hyderabad)

- Developed a Python-based metadata extraction pipeline for a Telugu literary corpus containing **758 books and 10,000+ stories (1947–2012)** using Gemini 2.5 Flash and structured JSON workflows.
- Built a Telugu **Retrieval-Augmented Generation (RAG)** system leveraging Qdrant and GTE Multilingual embeddings, achieving **92% Hit Rate@1, 99% Hit Rate@5, and MRR ~0.95** across **10,000+ stories**.
- Contributed to **3 production-grade AI applications** involving large-scale data processing, workflow automation, and backend system development for low-resource language technologies.

PROJECTS

- Telugu Agentic RAG System** — *Generative AI, Streamlit, Google Gemini OCR, RAG* [GitHub] [Live Link]
 - Built a **multi-agent Telugu RAG system** using Python, Qdrant, and GTE Multilingual embeddings over **10,000+ stories**, achieving **92% Hit Rate@1, 99%@5, and MRR ~0.95**.
 - Architected a **Planner → Drafter → Critic** workflow to iteratively refine generated stories, reducing stylistic drift and "Translated English" artifacts.
 - Engineered a **Validator Agent** enforcing **8 literary quality rules** and automated metadata extraction, embedding generation, and vector ingestion pipelines.
- Vidya Setu – Offline-Digital Learning Platform** — *React, Node.js, Capacitor, MongoDB, Socket.IO, FFmpeg* [GitHub]
 - Developed **Vidya Setu**, an edge-to-cloud learning platform using React, Node.js, Capacitor, and MongoDB for reliable education delivery in low-bandwidth environments.
 - Built offline synchronization pipelines with IndexedDB and Service Workers, ensuring **100% data survivability**, flushing **100 queued operations in 1.42 seconds**, and achieving a **0.67-second application startup time**.
 - Engineered FFmpeg-based video compression pipelines that reduced media payloads by **72.5%**, maintained peak memory utilization below **42 MB**, and delivered real-time interactions with **48 ms** latency.

TECHNICAL SKILLS

- Languages:** Python, JavaScript, C, HTML, CSS
- Frameworks & Technologies:** FastAPI, Express.js, React.js, Streamlit, REST APIs, Socket.IO
- Databases & Tools:** MySQL, MongoDB, Qdrant, Git, GitHub, Firebase, Cloudinary
- Libraries:** Pandas, NumPy, Scikit-learn, XGBoost, MediaPipe, Hugging Face
- GenAI & AI:** RAG, Multi-Agent Systems, LLM Applications, Vector Search, Prompt Engineering, Gemini/OpenAI APIs

LEADERSHIP

- Technical Lead & Treasurer, DevCatalyst Technical Club** – Organized hackathons and technical events, led club initiatives, and managed operational activities.

ACHIEVEMENTS AND CERTIFICATIONS

- Qualified for the **Final Round of Youth Speak Hackathon (IIT Hyderabad)**, participated in **4+ hackathons**, and attended **8+ technical events/workshops**.
- Published a **75K+ Telugu stories dataset** on Hugging Face and contributed to **3 production-grade AI systems** deployed at Viswam.ai.
- Certifications:** Saylor Academy Python Programming, Udemy Python Programming.