Shivam Thakur

Bengaluru, Karnataka | +91 9994728607 | shivamghe2@gmail.com | linkedin.com/in/shivam-thakur-swe | github.com/MavihsLH

Education

Manipal Institute of Technology (MAHE) — Manipal, Karnataka
M.Tech in Artificial Intelligence & Decision Science | **CGPA: 8.29 / 10.0** — *Aug. 2024 – May 2026*

Shanmugha Arts, Science, Technology & Research Academy (SASTRA) — Thanjavur, Tamil Nadu
B.Tech in Computer Science and Engineering — *Aug. 2020 – June 2024*

Experience

Data Science & AI Intern — July 2025 – May 2026
Aditya Auto Products and Engineering — *Bengaluru, Karnataka*

- Deployed a real-time acoustic motor fault detection system across shop-floor production lines, achieving **100%** fault recall at ∼**45ms** inference latency.
- Engineered an end-to-end audio DSP and explainable AI (Grad-CAM) pipeline in PyTorch using CNN and CRNN architectures to isolate mechanical bearing defects from noisy industrial audio.
- Deployed an on-premise enterprise LLM assistant (Llama 3, Qwen, Gemma) with custom RAG knowledge bases and built a full-stack web portal to digitize quality-assurance workflows.
- Engineered high-throughput barcode scanning and dynamic linking logic across **2** production Warehouse Management Systems, processing **100,000+** daily records to reconstruct multi-level assembly hierarchies.
- Developed a full-stack IoT telemetry platform using FastAPI, React, and Azure VMs, designing fault-tolerant data ingestion pipelines optimized for constrained factory networks.
- Built automated Python ETL pipelines with custom imputation logic to recover missing legacy sales data and architected production Power BI dashboards (DAX) to track real-time manufacturing efficiency.

Publications & Personal Projects

Vision-based Image Generation & Analysis Model (VIGAM) | IEEE Access — Published: Dec. 2024

- Co-authored an IEEE-published deep learning framework combining Real-ESRGAN, Vision Transformers (ViT), and GPT-2 to super-resolve aerial drone imagery for automated plant disease detection.
- Engineered a coupled Real-ESRGAN-ViT architecture achieving **90.0%** classification accuracy, outperforming standalone ViT baselines by **5.9%** and Swin Transformer variants by **15.4%**.

SILO: Local Agentic RAG Assistant | *Python, FastAPI, LangGraph, Qdrant, Ollama, PyTorch* Source Code

- Built a privacy-first Agentic RAG platform integrating FastAPI, Qdrant hybrid vector search (dense + sparse), and Ollama for low-latency local document indexing and real-time token streaming.
- Architected stateful LangGraph workflows executing Corrective RAG (CRAG) and Self-RAG loops with Cross-Encoder re-ranking and web search fallbacks to suppress hallucinations.

Algorithmic Portfolio Optimization Engine | *Python, Flask, CVXPY, NumPy, Pandas* Source Code

- Engineered a quantitative risk-minimization engine and Flask web application using convex optimization (CVXPY) to process **10+ years** of historical market data across **15+** equities.
- Formulated quadratic programming solvers based on Markowitz Modern Portfolio Theory to dynamically compute efficient frontiers, expected returns, and maximum Sharpe Ratio allocations.

Technical Skills

Languages: Python, C++, SQL, JavaScript, HTML/CSS

AI & Machine Learning: PyTorch, TensorFlow, Scikit-learn, XGBoost, Computer Vision (OpenCV, ViT, GANs), Audio DSP, XAI (Grad-CAM)

GenAI & LLMs: LangGraph, Agentic RAG (CRAG, Self-RAG), Ollama (Llama 3, Qwen, Gemma), Qdrant (Vector DB), SentenceTransformers

Backend, Cloud & Data: FastAPI, Flask, React, REST APIs, Docker, Linux, Git, Azure (VMs, IoT Hub), GCP (Vertex AI), ETL, Power BI (DAX)

Certifications: Google Cloud Skills Boost: Prompt Design in Vertex AI, Transformers & BERT, Responsible AI