SOHAM HARIDAS GAIKWAD

B.Tech CSE • SPIT Mumbai • 3rd Year (2024–2028)
soham092510@gmail.com • +91 8788156563 • Mumbai, India
GitHub • Portfolio • LeetCode

EDUCATION

B.Tech – Computer Science & Engineering *2024 – 2028*
Sardar Patel Institute of Technology (SPIT), Mumbai
CGPA: 8.09 | **Coursework:** Data Structures & Algorithms, Operating Systems, Computer Networks, DBMS, Computer Architecture, Data Science, Software Engineering, Theory of Computation

TECHNICAL SKILLS

Programming Languages: Python, C, Java

Machine Learning / Deep Learning / LLM: Machine Learning, Deep Learning (Neural Networks), Large Language Models (LLM), Retrieval-Augmented Generation (RAG), LangChain, Google Gemini API, FAISS (Vector Search), Prompt Engineering, Conversational Memory Design, Supervised Learning, Gradient Boosting, XGBoost, LightGBM, Hyperparameter Tuning, Predictive Analytics, Time Series Forecasting (Prophet), PyTorch, TensorFlow, Keras, Scikit-learn

Data Science & Analysis: Exploratory Data Analysis (EDA), Feature Engineering, Data Cleaning, Data Processing, Pandas, NumPy, Matplotlib, Seaborn, Jupyter Notebooks

NLP & AI Applications: Sentiment Analysis (VADER), Named Entity Recognition, Multilingual NLP (langdetect), Multimodal AI (Text + Image Reasoning), Semantic Search, Retrieval-Grounded Chatbots

Backend & Web Development: FastAPI (ASGI), WebSockets, REST APIs, SQLAlchemy ORM, Pydantic, Flask, Streamlit, HTML, CSS, JavaScript (ES6+), Chart.js, Web Scraping

Observability, Systems & Linux: CNCF Kepler (eBPF), Linux Kernel, Intel/AMD RAPL, Prometheus, Grafana, PromQL, Google OR-Tools (MILP), OpenCV, YOLOv8, Real-Time / Async Systems

Databases & DevOps: PostgreSQL, TimescaleDB, SQLite, MySQL, Database Management Systems (DBMS), Docker, Docker Compose, Git, GitHub Actions, CI/CD, Render Cloud, Linux, Windows, Conda/Environment Management

Domain & Fintech Engineering: Double-Entry Bookkeeping, ACID Transactions, SHA-256 Cryptography, Scope 2 Carbon Accounting, CEA India Grid Baseline, SEBI BRSR / ESG Compliance

CS Core: Data Structures & Algorithms, Operating Systems, Computer Networks, Computer Architecture, Software Engineering, Theory of Computation, Object-Oriented Programming (OOP)

UI / UX & Interpersonal: Figma (Basic), Wireframing, User-Centered Design, Problem Solving

PROJECTS

Real-Time Gen AI Customer Service Bot — 6-Feature LLM System *ElevanceSkills Internship | 2026*

- Built and extended a retrieval-augmented customer service chatbot (Google Gemini + LangChain + FAISS) into 6 production-style AI systems on top of a shared architecture.
- **Sentiment-Aware Chatbot:** integrated VADER sentiment classification to detect customer emotion and dynamically adapt response tone in real time.
- **Medical Q&A Chatbot:** built a custom XML parser to process 16,000+ MedQuAD Q&A pairs; implemented FAISS-based retrieval and keyword-driven medical entity recognition (symptoms/treatments).
- **Dynamic Knowledge Base:** designed an incremental vector-index update mechanism to add new information without full re-embedding.
- **arXiv Research Assistant & Multimodal Assistant:** built a live arXiv paper-search/summarization system with query-rewriting for follow-ups, plus a Gemini-powered text + image reasoning assistant with persistent conversational memory.
- **Multilingual Chatbot:** built automatic language detection and cross-lingual response generation (English, Hindi, Marathi, Spanish) with context retention across language switches.

Green-Finance — Real-Time Energy & Computational Carbon Observability for Financial Microservices *Personal Project | 2026*
GitHub • Live Demo | Tech Stack: Python, FastAPI, SQLAlchemy, CNCF Kepler (eBPF), Prometheus, Grafana, Docker, SQLite/PostgreSQL, Chart.js

- **Architected an Atomic Core Banking Engine:** Implemented double-entry bookkeeping (Σ Debits ≡ Σ Credits) with pessimistic row-level database locking and sub-45ms latency across simulated inter-bank UPI/NEFT transfers.
- **Microsecond Hardware Energy Profiling:** Integrated CNCF Kepler eBPF telemetry to profile CPU/DRAM energy draw via Linux Intel/AMD RAPL hardware registers, attributing ~0.695 Joules per atomic transfer.
- **Engineered 96.2% Attribution Accuracy Model:** Implemented an energy conservation validation model ($E_{host} = \Sigma E_{containers} + E_{idle} + \Delta E_{residual}$) with an Exponential Moving Average (EMA, α=0.2) filter, achieving **96.2% attribution fidelity** and ≤ 5.4% repeatability coefficient.
- **Statutory ESG Carbon Accounting & Cryptographic Proof:** Automated Scope 2 emissions calculations using Central Electricity Authority (CEA) of India baseline factors (713.0 gCO_2/kWh) and sealed every ledger transaction with immutable SHA-256 block hash chaining.
- **Full-Stack Enterprise Portals & Telemetry Pipelines:** Built three decoupled web interfaces (Core Banking Portal, Sustainability Investigator, and 7-panel Grafana v11 Prometheus dashboard) with automated SEBI BRSR / RBI compliance PDF and CSV report exporters.

Drive-Shield — Real-Time Edge ADAS & Incident Forensics Platform *Personal Project | 2026*

Tech Stack: Python, PyTorch, YOLOv8, OpenCV, FastAPI, WebSockets, SORT, SQLite, JavaScript

- Developed an edge-computing ADAS pipeline using YOLOv8 and SORT tracking, achieving 30 FPS inference (<45ms latency) for multi-class obstacle detection, monocular depth estimation, and optical flow speed tracking.
- Engineered predictive risk algorithms evaluating Time-to-Collision (TTC ≤ 2.0s), abrupt deceleration (>6.5 m/s²), and pedestrian fall/tumble postures to trigger millisecond safety warnings.
- Built a FastAPI & WebSocket dual-stream cockpit UI with an in-memory rolling buffer, automatically capturing and archiving 30-second pre/post-incident forensic video evidence signed with SHA-256 cryptographic hashes.

BIT-SHIELD — Bitcoin Forensics & Anomaly Detection Engine *Personal Project | 2026*

Tech Stack: Python, FastAPI, NetworkX, Scikit-Learn, React 18, TypeScript, Cytoscape.js, SQLite

- Engineered an air-gapped, zero-cloud Bitcoin transaction forensics platform featuring a universal offline data fabric with fuzzy regex schema normalization for arbitrary CSV/JSON datasets and automated SHA-256 chain-of-custody verification.
- Implemented an algorithmic graph mining engine detecting 8 behavioral money-laundering motifs (Rapid-Hop, CoinJoin mixing, Peel Chains, High Fan-Out/In) across directed bipartite transaction topologies (Address ↔ Transaction).
- Built a 5-lens risk fusion pipeline fusing deterministic rules (R1–R8) with unsupervised Isolation Forest scoring, generating explainable, court-admissible evidence chains.
- Designed an interactive SOC forensic workstation with Cytoscape.js supporting 4 dynamic graph layouts, live entity search, and 2x high-resolution PNG evidence export.

LearnLens AI — Multimodal Lecture Observation & Knowledge Synthesis Agent *Personal Project | 2026*

GitHub • Live Demo • Backend API • API Docs | Tech Stack: React 19, TypeScript, FastAPI, WebSockets, Python, Faster-Whisper, OpenCV, Ollama/LLMs, ReportLab, python-pptx

- Architected a real-time multimodal learning agent that observes user-authorized educational Chrome tabs (Coursera, YouTube, Udemy) and synthesizes grounded lecture slide decks, notes, and interactive quizzes.
- Engineered a dual client-server audio transcription pipeline using browser Web Speech STT and containerized Faster-Whisper (int8, VAD filtering), streaming timestamped transcripts over full-duplex WebSockets.
- Built a perceptual frame-extraction engine combining client-side difference hashing with OpenCV Hamming-distance comparison, cutting redundant frame transfer by 85% while capturing genuine slide transitions.
- Developed automated study-export generators with ReportLab and python-pptx, producing 16:9 slide decks, OCR annotations, and study packs with single-click download.
- Deployed full-stack cloud infrastructure via Render Blueprints, tuning SQLite (WAL mode, B-tree indices) for sub-millisecond query latency across 500+ session records.

GridPulse (NeighbourFlex) — AI Energy Intelligence Platform *Schneider Electric Challenge | 2026*

GitHub • Live Demo • Backend API • API Docs • CI/CD Pipeline | Tech Stack: Python (FastAPI), TypeScript, React, Google OR-Tools, PostgreSQL/TimescaleDB, Docker, WebSockets

- Built a real-time IoT energy-intelligence platform for the Schneider Electric Innovation Challenge (PS3) to ingest, validate, and optimize distributed energy assets (solar PV, EV fleets, 100 kWh BESS) across a 100-endpoint microgrid.
- Formulated a Mixed-Integer Linear Programming (MILP) model with Google OR-Tools to optimize multi-period battery dispatch and EV charge deferral, achieving 101.4% planned-vs-measured dispatch efficiency.
- Engineered an autoregressive multi-horizon forecasting engine (15m/1h/6h/24h) integrating live weather telemetry, achieving a 96.2% confidence score and 2.4 kW MAE (3.8% MAPE).
- Implemented a stratified baseline engine and multi-factor anomaly diagnostics that attribute demand spikes to heatwaves versus synchronized EV charging in real time.
- Designed an asynchronous FastAPI + SQLAlchemy (async) + WebSocket backend with sub-2s telemetry streaming, SlowAPI rate-limiting, Docker Compose packaging, and automated GitHub Actions CI/CD.

RAG-Based AI Teaching Assistant *Personal Project | 2025*

- Built a Retrieval-Augmented Generation pipeline combining an LLM with vector-based retrieval to answer course questions grounded in uploaded study material, reducing hallucination and citing sources.

HyperLocal Demand Forecaster V2 *Data Science Course Project | 2025*

- Ensemble model (LightGBM + XGBoost + CatBoost) with Optuna hyperparameter tuning and Ridge meta-learner stacking.
- Built a SHAP explainability layer and Indian-festival-proximity features to capture local demand patterns.

CERTIFICATIONS & COURSES

Real-Time Gen AI Project Internship — Customer Service Bot *ElevanceSkills | Jul 2026*

- Mentor-guided internship extending a base LLM chatbot into 6 production-style AI systems (sentiment analysis, medical Q&A, dynamic knowledge base, research assistant, multimodal AI, multilingual support).
- Training Certificate: View Certificate • Internship Completion Certificate: View Certificate

Data Science with Python *CodeWithHarry | Mar 2026*

- Covered end-to-end data science workflows — EDA, ML model building with Scikit-learn, and reproducible research; hands-on with Pandas, NumPy, Matplotlib, neural networks, and LLM concepts. Certificate: View Certificate

ACADEMIC HIGHLIGHTS

- **CGPA 8.09** in a competitive CSE program at SPIT — consistently pushing for higher scores.
- Strong in core subjects: Operating Systems, Computer Networks (subnetting, TCP/IP, routing), DAA, DBMS, Software Engineering, Theory of Computation.