

SOHAM GORE

Mumbai, Maharashtra

☎ +91 8591485946

✉ sohamrgore@gmail.com

🌐 [Website](#)

🌐 [LinkedIn](#)

🐙 [GitHub](#)

Education

KJ Somaiya School of Engineering

2024 – Expected 2028

Bachelor of Technology in Electronics & Computer Engineering; CGPA: 9.19

Mumbai, Maharashtra

Honours in Data Science & Analytics

Experience

FlyRank AI

July 2026 – August 2026

Machine Learning Engineering Intern

Remote

- Formulated a search content refresh scoring pipeline across 30,000 performance records (32 client sites), stripping temporal trend fields to enforce strict data leakage controls.
- Evaluated classification architectures under a grouped client holdout split (8 unseen clients, 7,115 test rows) to benchmark model generalization against a heuristic baseline.
- Boosted review queue precision to **0.80 P@50** (vs. 0.52 baseline) and **0.85 P@20** across held-out client domains.
- Engineered a reason-coded decision engine, deploying a reproducible end-to-end ML research paper and action playbook.

KJ Somaiya School of Engineering

July 2026 – Present

Research Intern, Advisor: Dr. Ayesha Hakim

Mumbai, Maharashtra

- Proposed and finalized the end-to-end architecture for a hybrid multimodal fake news detection framework integrating LLMs and Graph Neural Networks for evidence-first reasoning.
- Developing the system's core components, including dual-modality (text-image) encoders, a FAISS-backed RAG verification module, and a Heterogeneous Graph Attention Network (PyTorch Geometric) for cross-modal and propagation analysis.

Elevate Labs

June 2025 – July 2025

AI/ML Intern

Remote

- Engineered a deep Convolutional Neural Network (CNN) in Keras using Mel-Spectrogram audio representations, achieving **89% test accuracy** across 10 distinct musical genres on 1,000+ tracks.
- Developed an automated spectral signal processing pipeline via Librosa to extract 20+ Mel-Frequency Cepstral Coefficients (MFCCs), spectral centroids, and zero-crossing rates, accelerating model training convergence by **25%**.
- Analyzed class confusion matrices via custom diagnostics, boosting overall model F1-score by **12%**.
- Built a low-latency inference engine for the `.keras` artifact, enabling sub-**200ms** audio predictions.

Selected Projects

SAR Deforestation Monitor | Python, Google Earth Engine (GEE), XGBoost, OpenCV | [Git](#)

May 2026

- Engineered a GEE pipeline ingesting Sentinel-1 SAR imagery to enable all-weather forest change monitoring.
- Developed a Rust/PyO3 spatial filtering kernel, achieving a **10x speedup** over Python for SAR speckle reduction.
- Trained XGBoost on SAR texture metrics and VV/VH ratios, achieving **95% accuracy** across 52,000+ spatial pixels.
- Orchestrated a YAML-configured pipeline unifying satellite ingestion, PyO3 transforms, and XGBoost inference.

Hybrid Dynamic Pricing & Valuation Engine | Python, LightGBM, SHAP, FastAPI, Streamlit | [Git](#)

January 2026

- Engineered a LightGBM pipeline with TF-IDF features, achieving a **2.96% MAPE** across 1,200 listings.
- Integrated SHAP tree explainers to compute feature-level value attributions for transparent price breakdowns.
- Deployed a low-latency FastAPI REST service delivering real-time dynamic price multiplier inference.
- Built an interactive Streamlit dashboard to simulate real-demand elasticity and scenario-based price adjustments.

Technical Skills

Programming & Development: Python, Java, SQL, FastAPI, Docker, Git

Geospatial & Remote Sensing: Google Earth Engine, Sentinel-1 SAR, OpenCV, Rust/PyO3

ML, DL & GenAI: TensorFlow, PyTorch, PyTorch Geometric, NLP, XGBoost, CNNs, GNNs, RAG, FAISS

Data Pipelines & Preprocessing: EDA, Feature Engineering, Feature Selection, PCA, Dataset Management

Production & Deployment: CI/CD Pipelines, AWS, Streamlit, Flask

Leadership / Extracurricular

Datazen Somaiya

July 2025 – July 2026

Technical Team Member

KJ Somaiya School of Engineering

- Spearheaded curriculum execution and capstone mentoring for Data Trek, training **350+ participants** in ML.
- Orchestrated technical problem statements and evaluation pipelines for Datathon 2026, reaching **1,600+ registrants**.

Open Source Contributions

etsi-ai

Aug 2025 – Present

Maintainer & Core Developer | [Git](#)

Open Source Organization

- Launched **etsi-etna** on PyPI, optimizing a neural network framework combining Rust speed and Python runtimes.
- Engineered drift detection capabilities within **etsi-watchdog** to automate root-cause analysis of model failures.
- Managed architecture as a maintainer, establishing strict automated validation pipelines and release roadmaps.