

Saurav Kumar

linkedin.com/in/sauruv • github.com/sauruv • huggingface.co/sauruv • sauruv.github.io

EDUCATION

Meghnad Saha Institute of Technology | Kolkata, WB

Oct. 2022 – Jul. 2026

B.Tech in Computer Science and Engineering, Specialization in Artificial Intelligence and Machine Learning

Relevant Courses: Data Structures and Algorithms, Object-Oriented Programming, Operating Systems, DBMS, Computer Networks, Artificial Intelligence, Machine Learning, Deep Learning, Probability and Statistics, Computer Vision

TECHNICAL SKILLS

Languages: Python, TypeScript, SQL

Generative AI: LLM APIs, Retrieval-Augmented Generation (RAG), Embeddings, Retrieval and Reranking, Structured Outputs, Tool Calling, AI Agents, LLM Evaluation and Observability, Multimodal AI

AI/ML: PyTorch, YOLO, OpenCV

Engineering Stack: React, Next.js, FastAPI, PostgreSQL, pgvector, Redis, Celery, Docker, Vercel, AWS, GitHub Actions

RELEVANT EXPERIENCE

Dentcare Dental Clinic | Kolkata, WB

Apr. 2026 – May 2026

Computer Vision Engineer

- Designed and delivered an end-to-end computer vision prototype for binary classification of oral lesions as benign or malignant, combining fine-tuned DINOv2 and EfficientNetV2-B3 models in an ensemble architecture
- Achieved a test set AUC-ROC of approximately 0.96 and integrated Grad-CAM-based visual explanations to improve model interpretability and support qualitative review of predictions
- Packaged the ensemble as an offline PyQt6 application and exported it to ONNX, enabling portable inference on clinic hardware on a four-week prototype engagement

PROJECTS

Amaya | Multimodal Deepfake Detector | [Demo](#)

Jul. 2026

- Developed a multimodal deepfake video detector using SigLIP2, rPPG, CNN-Transformer encoding, and quality-aware cross-attention fusion to combine visual and physiological clues across multiple temporal windows
- Built leakage-aware training and zero-shot evaluation pipelines across CelebDF v2, FaceForensics++ C23, and WildDeepfake, achieving 95.73% in-domain accuracy and 79.10% cross-dataset macro accuracy

Ukti | Speech Transcription and Translation | [Demo](#)

Aug. 2026

- Developed a browser-only speech transcription and English translation application using Whisper, Transformers.js, ONNX Runtime Web, and Silero VAD with local file and microphone input
- Built capability-adaptive WebGPU/WASM inference, WebWorker orchestration, bounded-memory audio processing, timestamped editable transcripts, and offline caching for privacy-preserving client-side execution.

ACHIEVEMENTS

SAP Hackfest 2025 College Round Winner

May 2025

SIH 2025 Internal Hackathon Winner

Sep. 2025