AYUSH SAUN

ayushsaun@gmail.com | (+91) 9667629472 | ayushsaun | ayushsaun24024 | Portfolio

Professional Summary

Engineer & Applied Researcher with **2+ years** turning data into production value through cloud-native **ETL** pipelines, big-data speech processing and robust **data-modeling**. Shipped **40+** zero-downtime releases at Samsung—launching a green-field portal in < 6 months and sustaining **99.8%** uptime—before halving audio-deepfake error to **EER 6.3%** on **575 h** of speech at IIIT-Delhi with reproducible **PyTorch** MLOps that ran experiments **3×** faster. Owns the full lifecycle—data ingestion, feature engineering, model training, CI/CD, and observability—using Python, SQL, AWS, Docker, and Git to deliver measurable, production-ready impact.

Education

M.Tech Computer Science (CGPA: 7.69) — **IIIT-Delhi**
Aug 2024 - Present — Delhi, India

B.Tech Electrical Engineering (CGPA: 7.84) — **Delhi Technological University**
Aug 2018 - June 2022 — Delhi, India

Experience

Post-Graduate Researcher — **Jan 2025 – Present**
Infosys Centre for Artificial Intelligence, IIIT-Delhi — Delhi, India

- Designed a **modular ETL/train/validate/infer pipeline** for speaker verification and anti-spoofing, processed **500+ h** of speech, and cut pipeline build time **3×**.
- Fine-tuned SOTA **self-supervised encoders** (HuBERT, Wav2Vec2, WavLM) with CNN heads (AASIST, ECAPA-TDNN), slashing spoof **EER to 6.3%** (-50%).
- Applied **RawBoost**, **MUSAN**, and **RIR** augmentation, reducing tandem-EER to **30%** and boosting accuracy **+12%**.
- Led end-to-end research ops data curation, hyper-parameter sweeps, ablations, GPU scheduling—and benchmark models.
- **Tech Stack: Python**, **PyTorch**, **TorchAudio**, **HuggingFace Transformers**, **HuBERT**, **Wav2Vec2**, **WavLM**, **TitaNet**, **ECAPA-TDNN**, **RawBoost**, **MUSAN**, **RIR**, **Git/GitHub**, **Linux**, **CUDA**

Software Engineer — **Jun 2022 – Jul 2024**
Samsung R&D Institute India – Delhi — Delhi, India

- Architected & maintained **4** internal portals (∼100 users) as a full-stack developer; shipped REST APIs, tuned databases, and automated **CI/CD** for **40+** releases/yr with **99.8%** uptime.
- Engineered responsive UIs with **React.js** & **Material-UI**; added **Jest** tests that lifted engagement **25%** and cut load times **35%**.
- Developed secure **Spring Boot** micro-services with token-based authentication, **JUnit**, and **Swagger**, accelerating developer velocity.
- Optimized **Oracle SQL** via caching & query tuning, reducing API latency **40%** and trimming infrastructure cost.
- Leveraged **AWS** (**S3**, **EC2**, **IAM**, **CloudFront**, **CloudWatch**) for compliant, monitored infrastructure.
- Orchestrated Agile sprints with **Jira** & **Confluence**, boosting team velocity **20%**.
- Spearheaded a hackathon prototype using **Angular.js**, **Python**, and **Hugging Face** ML for advanced analytics.
- Drove end-to-end delivery of the 4^{th} portal—requirements, UX collaboration, QA coordination, deployment, and post-release support—launched in $<$ **6 months**.
- **Tech Stack: React.js**, **Material-UI**, **Angular.js**, **Jest**, **Java Spring Boot**, **Python**, **Swagger**, **JUnit**, **Oracle SQL**, **AWS** (S3, EC2, IAM, CloudFront, CloudWatch), **Git/GitHub**, **CI/CD**, **Jira**, **Confluence**, **Jenkins**

Projects

Audio Deepfake Detection — **Jan 2025 – Present**

- Evaluated SOTA anti-spoofing models (**WavLM Base**, **ECAPA-TDNN**, **RawNet2**) on **575 h** of speech using a modular **PyTorch** k-fold pipeline, securing statistically robust metrics.
- Adapted SSL encoders (**HuBERT**, **Wav2Vec2**, **WavLM**) with CNN heads (**AASIST**, **ECAPA-TDNN**), halving spoof **EER** to **6.3%**.
- Optimized training through Bayesian hyper-parameter search and waveform augmentations (**RawBoost**, **MUSAN**, **RIR**), cutting iteration time **30%**.
- Integrated **SpeechBrain** and **HuggingFace Transformers** for one-click dataset/model loading, checkpointing, and experiment tracking.
- Oversaw end-to-end research workflow—dataset curation, GPU scheduling, ablation studies, and documentation—benchmarking results.
- **Tech Stack: Python**, **PyTorch**, **SpeechBrain**, **HuggingFace Transformers**, **HuBERT**, **Wav2Vec2**, **WavLM**, **TitaNet**, **ECAPA-TDNN**, **RawNet2**, **RawBoost**, **MUSAN**, **RIR**, **CUDA**, **Git/GitHub**, **Linux**

Automatic Speaker Verification System **Jan 2025 – Present**

- Built an end-to-end biometric speaker-verification pipeline with **deep speaker embeddings**, **residual-phase features**, and **ensemble score fusion**; trained on **350 h** multilingual speech for secure enterprise sign-in.
- Tuned similarity scorers (**cosine**, **PLDA**) and feature parameters, trimming tandem **EER to 30%**.
- Owned the full delivery cycle—data prep, GPU scheduling, hyper-parameter sweeps, evaluation, deployment, and post-release support—meeting all release deadlines.
- **Tech Stack: Python**, **PyTorch**, **SpeechBrain**, **HuggingFace Transformers**, **CUDA**, **Git/GitHub**, **Docker**, **Linux**

ScholarAI – AI-Powered PDF Learning Assistant **Mar 2025 – Jun 2025**

- Developed an AI-first web app that transforms academic PDFs into structured, interactive learning formats—**summaries**, **study notes**, **flashcards**, and **quizzes**—to accelerate student comprehension and retention.
- Designed and implemented end-to-end document processing workflows using **Genkit** and **LangChain**, leveraging **Gemini Pro** and **GPT-4 Turbo** for structured content generation aligned with UX schemas.
- Architected dynamic UI using **React (Next.js 15)** and **Tailwind CSS**, including accordion-based summaries, editable note blocks, click-to-flip flashcards, and real-time scored quizzes with explanations.
- Engineered document parsing and ingestion using **pdfjs-dist** and **pdfplumber**, enabling seamless processing of multi-page, text-based academic content.
- Integrated with **Firebase Hosting**, **Firestore**, and **Storage** for quiz state persistence, document uploads, and downloadable assets (PDFs, CSVs, JSONs).
- Deployed full-stack application to **Vercel** with production-grade build optimizations via Turbopack and custom LLM execution flows using `tsx` and hot-reload dev mode.
- **Tech Stack: TypeScript**, **Next.js 15**, **React 18**, **Tailwind CSS**, **Genkit**, **LangChain**, **GPT-4 Turbo**, **Gemini Pro**, **Vercel**

Single-Object Tracking System **Aug 2024 – Dec 2024**

- Delivered a real-time tracker with **camera-motion compensation** via ORB-based affine alignment, sustaining **30 FPS** on 1080p streams.
- Integrated multi-scale search windows and hybrid descriptors (**HOG**, **LBP**, **SIFT**, **ORB**), strengthening robustness to shape, texture, and scale variation.
- Employed ensemble regressors (**Linear Regression**, **Random Forest**) for centroid/box prediction, achieving **85% IoU**, $R^2 = 0.92$, and cutting **MAE 20%**.
- Owned the full delivery cycle—dataset curation, hyper-parameter sweeps, performance profiling, Git-based code reviews, deployment, and post-release support—integrating the module into downstream CV pipelines.
- **Tech Stack: Python**, **OpenCV**, **HOG**, **LBP**, **SIFT**, **ORB**, **Linear Regression**, **Random Forest**, **NumPy**, **SciPy**, **Matplotlib**, **Git**, **Linux**

Skills

Programming Languages & Scripting: Python, SQL, C++, Java, JavaScript, TypeScript

AI, Machine Learning & Data Science: PyTorch, TensorFlow, scikit-learn, Hugging Face **Transformers**, ETL Pipelines, Data Modeling, Hadoop, SpeechBrain, torchaudio, OpenCV, librosa, Pandas, NumPy, Matplotlib, CUDA

Cloud & DevOps: AWS (S3, EC2, IAM, CloudFront, CloudWatch), Docker, Jenkins, Git/GitHub, Linux, RESTful APIs, CI/CD, Jira, Confluence

Web & Software Development: React, Angular, Material-UI, Java Spring Boot, Swagger/OpenAPI, Jest, JUnit, Full-Stack Development

Productivity Tools & IDEs: VS Code, Jupyter Notebook, Google Colab, Postman, LaTeX/Overleaf, Anaconda, IntelliJ IDEA

Professional Skills: Research Development, Agile/Scrum, Technical Documentation, Problem-Solving, Project Management, Cross-Functional Collaboration, Release Coordination, Presentation Skills

Interests & Achievements

Technical Highlights:

- Completed **200 + competitive-programming questions**, strengthening data-structures, algorithms, and analytical reasoning.
- Passed **Samsung R&D Code Competency Test**, ranking in the top performance band for coding efficiency and software design.
- Achieved **runner-up** position among **50+ teams** in a company hackathon by **rapidly developing a AI/ML based solution** under **tight deadlines**.

Certifications

- **Programming with Python: Hands-On Introduction for Beginners** — Udemy **2020**
- **Front-End Web Development with React** — Coursera **2020**
- **Front-End Web UI Frameworks and Tools: Bootstrap 4** — Coursera **2020**