SHUFAN SHAHI

Dhaka, Bangladesh
+880 1990-935044 shufanshahi@gmail.com
github.com/shufanshahi linkedin.com/in/shahi-shufan shufanshahi.github.io/eng

About

AI/ML engineer and CSE undergraduate (IUT, 2026) who takes models from data analysis and experimentation to production APIs. Currently an AI Engineer at GLiTch LABS building, evaluating and deploying ML and LLM systems like RAG. Previously worked on fine-tuned ML model integration at Battery Low Interactive, built computer vision and tracking pipelines for drones at Project Aero and, as a remote ML Researcher at Elite Research Lab (New York), turned cutting-edge research into working, evaluated implementations.

Education

Islamic University of Technology **2022–2026**
B.Sc. in Computer Science and Engineering **CGPA: 3.58**
Coursework: Algorithms, Databases, Networks, Machine Learning and Deep Learning, Calculus, Probability and Statistics, Linear Algebra.

Chuadanga Govt. College **2021**
Higher Secondary Certificate (HSC) **GPA: 5.00/5.00**

Experience

AI Engineer **Oct 2026–Present**
GLiTch LABS *Dhaka, Bangladesh*

- Developed full-stack web applications and backend services with REST APIs, database schemas, and responsive frontends.
- Architected a RAG pipeline (lexical and vector retrieval with reranking) and tuned an LLM, improving answer accuracy and reducing hallucinated responses.
- Delivered a real-time voice agent (streaming speech-to-text, LLM, text-to-speech) and significantly reduced end-to-end response latency through WebSocket streaming, sentence-level TTS chunking, and response caching.

Researcher **Aug 2026–Present 2026**
Elite Research Lab *New York, USA (Remote)*

- Worked as an ML engineer on cutting-edge research projects, implementing recent methods, running experiments and evaluating model performance.

Intern **Oct 2025–Oct 2025**
Battery Low Interactive Ltd. *Dhaka, Bangladesh*

- Implemented Flutter app features and integrated 12 REST API endpoints, including a locally fine-tuned Qwen2.5-1.5B model.

Senior Executive **Feb 2026–Aug 2026**
Project Aero *Dhaka, Bangladesh*

- Led a team of 8 engineers creating software for the university drone program and international UAV competitions.

Researcher **May 2025–Sept 2025**
IUT Computer Vision Lab *Dhaka, Bangladesh*

- Researched multimodal AI methods to improve integration and mitigate modality imbalance.

Projects

DocuMind - Document Q&A | *Python, LangChain, LlamaIndex, ChromaDB, Ollama* GitHub
- Built a local RAG system for documents with cited answers and hallucination guardrails.
- Compared 4 chunking methods and 2 embedding models; added hybrid BM25 and reranking.

Benhallu - Bengali Hallucination Detection | *Python, RAG, BanglaBERT, scikit-learn* GitHub
- Built a retrieval-augmented pipeline to verify Bengali answers against supplied context or a Bengali QA knowledge base.
- Created a hallucination dataset and used fine-tuned BanglaBERT with answer-comparison features to classify responses.

Bengali Speaker Diarization - Speaker Turn Detection | *PyTorch, Silero VAD, ECAPA-TDNN* GitHub
- Built a speech pipeline using neural voice activity detection, speaker embeddings and agglomerative clustering to label timestamped speaker turns.
- Estimated speaker counts with silhouette scores and exported diarization results to CSV.

Drone Vision - Aerial Tracking | *PyTorch, YOLOv11X, ByteTrack* GitHub
- Fine-tuned YOLOv11X on VisDrone and used ByteTrack for human counting and multi-object tracking.

Biztrack - Business Analytics | *Next.js, Node.js, PostgreSQL* GitHub Demo
- Built full-stack AI modules for forecasting, customer segmentation and risk alerts on transactional data, backed by PostgreSQL and served through APIs.

Postfolio - Professional Networking Platform | *Spring Boot, Next.js, Docker, RabbitMQ, Stripe* GitHub Demo
- Built a containerized full-stack platform with Spring Boot backend services, RabbitMQ messaging and Stripe payments.
- Built an AI matching engine that ranks candidates against job postings to surface the best-fit hires, alongside portfolio and CV creation, job search and AI interview preparation.

Redrick Routinson - Campus Automation | *Python, React, Flask, Firebase* GitHub Demo
- Built genetic-algorithm scheduling optimization, QR exam attendance, class swapping, events, forums and lost-and-found.

Technical Skills

Programming Languages: Python, SQL, C++, Java
Machine Learning and Data: PyTorch, TensorFlow, scikit-learn, Pandas, NumPy
Generative AI and NLP: RAG, LLMs, LangChain
Computer Vision and Audio: YOLOv11, ByteTrack
Backend and APIs: REST APIs, Flask, Spring Boot, Node.js, RabbitMQ
Databases: PostgreSQL, MySQL, Firebase, ChromaDB
DevOps and Tools: Docker, Git, Linux
Web and Mobile: React, Next.js, React Native, Flutter

Achievements and Awards

- **5th Place, VAND4 International Competition** — Zero-Shot Industrial Track **2026**
- **First Runner-up, IUT Coderush Hackathon** — Rapid team application development **2025**
- **7th Place, BAIC Datathon** — Data preprocessing, visualization and modeling **2025**