

SHOUMIK SARKAR

Dhaka, Bangladesh | +8801776892842 | ssavi.pro03@gmail.com
linkedin.com/in/shoumik-sarkar-nsu | github.com/ssavi45

PROFESSIONAL SUMMARY

AI/ML Engineer focused on multimodal learning, computer vision, and building end-to-end digital products. Experienced in translating complex datasets into actionable insights, designing and deploying scalable AI tools, and collaborating across cross-functional workflows to deliver responsible, impact-oriented technology solutions. Proficient in Python with strong foundations in OOP, data structures, algorithms, and software engineering best practices.

EDUCATION & CERTIFICATIONS

- North South University, Dhaka - BSc in Computer Science and Engineering** May, 2026
- Specialization: Intelligent System Engineering
 - CGPA: 3.2/4.0
 - Relevant Coursework:** Data Structures & Algorithms, Theory of Computation, Programming Languages (C, C++, Java), Machine Learning, Deep Learning, Pattern Recognition & Neural Networks, Computer Vision
- OpenCV University - Vision Language Models (VLM) Bootcamp** January, 2026
- Comprehensive training on Multimodal AI architectures and Vision Language Models

PROFESSIONAL EXPERIENCE

- Flyrank AI - Machine Learning Engineer (Intern)** Remote | July, 2026 – Present
- Contributing to the design and development of end-to-end ML pipelines and AI-driven digital products, supporting process efficiency and responsible deployment.
 - Working alongside senior engineers to define system requirements and translate them into technical solutions for local, scalable AI tools.
 - Assisting in dataset evaluation and performance analysis using Python-based data processing workflows to support iterative model improvement.

PROJECTS AND RESEARCH

- DeceptiLens: Multimodal Deception Detection | 2025 – 2026** 
- Tech Stack: PyTorch, OpenCV, Librosa, MediaPipe, FastAPI, Streamlit*
- Built an end-to-end multimodal data pipeline using Python (OOP) combining facial landmarks, audio descriptors, and deep visual embeddings.
 - Extracted facial blendshape features via MediaPipe, audio features via Librosa, and visual embeddings via ResNet-18 for a unified data representation.
 - Designed a custom cross-attention architecture in PyTorch for bidirectional audio-visual feature fusion, achieving 82.11% accuracy on the DOLOS dataset.
 - Developed and deployed the system as an interactive Streamlit app; debugged and optimized inference pipeline with real-time data visualizations for performance metrics.
- RumorFusionX: Multimodal Fake Social Media Post Detection | 2025** 
- Tech Stack: PyTorch, OCR, Transformers, LLMs, OpenCV, FastAPI*
- Developed a multimodal misinformation detection pipeline combining OCR-extracted text, image features, and LLM-generated contextual descriptions.
 - Collected, cleaned, and analyzed a multilingual quantitative dataset of 2,500 social media images, applying supervised learning techniques to translate findings into structured training inputs.
 - Implemented advanced feature fusion techniques via REST API-backed services, achieving 93.07% classification accuracy over unimodal baselines.

MedHelp: Medical Information Assistant | 2024 – 2025



Tech Stack: Transformers, FastAPI, PostgreSQL

- Participated in the scoping and backend build of a conversational AI tool utilizing retrieval-augmented generation (RAG) to provide data-driven responses.
- Processed structured medical reference datasets to improve response relevance and implemented medicine price aggregation for comparative search.

Football Genius: Full Stack Web Application | 2026



Tech Stack: HTML5, CSS3, JavaScript, Supabase, Google OAuth, Gemini AI API, GitHub Actions, Node.js

- Built and deployed a multi-game-mode football trivia platform end-to-end with mobile-first responsive design, semantic HTML5, custom CSS design systems, micro-animations, and a full dark-mode UI.
- Implemented a complete user authentication system supporting email/password, Google OAuth, and OTP verification via Supabase Auth, including a user dashboard, score history, and profile management.
- Automated live player statistics updates using a GitHub Actions CI/CD pipeline with a scheduled Node.js scraper, syncing data to Supabase without manual intervention.
- Integrated Supabase JS SDK, Google OAuth, and Google Gemini AI API into a cohesive production architecture.

SKILLS SUMMARY

Programming Languages: Python, C++, JavaScript, SQL

Machine Learning & AI: PyTorch, TensorFlow, Scikit-learn, Transformers, Computer Vision, OpenCV, Hugging Face, Multimodal Learning, NLP, LLMs, Supervised & Unsupervised Learning, LLM Fine-Tuning

Backend & Data Processing: FastAPI, PostgreSQL, Supabase, REST APIs, Pandas, NumPy, MySQL, Data Pipelines, Data Cleansing

Tools & Operations: Git, GitHub Actions, VS Code, Linux, MS Office, Dashboards & Data Visualization, OOP, Software Engineering Best Practices

Core Competencies: Data Structures & Algorithms, Product Management, Research & Documentation, Cross-functional Collaboration, Analytical Problem Solving

EXTRACURRICULAR ACTIVITIES

NSU Computer and Engineering Club (NSUCEC) - In-charge of Team Logistics and Operations

2022 – 2023

- Managed end-to-end operational plans, procurement workflows, and resource allocation for major university technical events like [Bit Arena 2022](#) and many other workshops.
- Coordinated project timelines and cross-functional task distribution across student teams, identifying and resolving logistical bottlenecks to ensure process efficiency.
- Maintained internal tracking dashboards and documentation to streamline communications and maintain strict adherence to event timelines.