Aayaan Sahu

(669)-294-0189 | aasahu2@illinois.edu | linkedin.com/in/aayaan-sahu | github.com/Aayaan-Sahu

Education

University of Illinois Urbana-Champaign Champaign, IL

B.S. Computer Science + Economics *Expected Graduation: May 2028*

GPA: 4.0

Classes: Data Structures and Algorithms, Linear Algebra, Discrete Structures, Statistical Analysis

Clubs: Hack4Impact, Illinois Space Society

Skills

Languages: Python, R, SQL, TypeScript, C++, Java, Go
Frameworks: React, Django, Flask, Next.js, Express, FastAPI
Libraries/Tools: TensorFlow, PyTorch, scikit-learn, OpenCV
Systems: Git, Docker, AWS, PostgreSQL, MongoDB, Vector Databases
Concepts: REST Api, NLP, RAG, LLM, Computer Vision, Model Interpretability

Experiences

Operation Safe Escape | *Empowering Domestic Abuse Survivors* Sep 2025 – Dec 2025

- Built a dynamic form-builder module using **React** and **Django** allowing admins to author and deploy survivor-intake workflows, enabling secure data capture across **1000+** intake submissions
- Developed a real-time messaging system using **WebSocket** + **end-to-end encryption** that connects survivors and staff, maintaining secure audit logs
- Architected secure **RESTful APIs** to centralize logic across survivor intake and messaging, ensuring consistent data validation and traceable request auditing

Researcher — Analyzing Bias in Emergency Room Triaging Jun 2021 - Dec 2022

Aspiring Scholars Directed Research Program *Fremont, CA*

- Preprocessed 10+ large scale **US medical system** ER datasets (> 1M rows) to predict triage outcomes
- Trained and evaluated Random Forest, XGBoost, and neural networks to predict triage outcomes
- Implemented **SHAP values for interpretability**, identifying statistically significant model bias factors
- Published in **Annals of Biomedical Science and Engineering** and Stanford's JUST Health (pg. 13)
- Presented findings at Southern California Conference for Undergraduate Research (11/19/22)

Projects

LLMBDA | *Prompt Optimization Tool* Nov 2025 – Present

- Architected a full-stack SaaS that optimizes LLM prompts to reduce API overhead, achieving **10-20x** token compression
- Engineered a custom task routing algorithm to intelligently preprocess and direct prompts to optimized LLM configurations, reducing latency and enhancing output accuracy and efficiency
- Developed a production-ready web platform using React with integrated **Stripe** payments and subscription management, managing the end-to-end lifecycle from development to a successful **Chrome Web Store launch**

SNAPSHOT | *Full Stack Mobile App* Dec 2025 - Present

- Engineered a scalable **Go backend** with **RESTful APIs** and **PostgreSQL** to handle authenticated users, group-based photo submissions, and time-windowed constraints, ensuring consistent data integrity across concurrent clients
- Designed a cross-platform **React Native** app using Axios-based API integration, enabling real-time photo uploads, group feeds, and client-side validation with low-latency backend communication
- Integrated **AWS S3** with presigned URLs for secure, direct photo uploads, reducing backend load

TalkPDF | *Retrieval-Augmented Generation System* Jun 2025 - Aug 2025

- Engineered a full-stack web app integrating Python, **SQL**, Supabase, Stripe, **Ollama**, and **vector databases**
- Implemented document ingestion, cleaning, and querying to allow users to "chat with a PDF"
- Designed clean **React** frontend with ShadCDN and ensured seamless UX across browsers