

ANSHUL SHAH

📞 972-948-7760 ✉ shahanshul@gmail.com 🔗 linkedin.com/in/anshul-shah2 🐙 github.com/shahanshul8052 🌐 portfolio

EDUCATION

Purdue University: Bachelor of Science in Computer Science

West Lafayette, IN

Concentration in Machine Learning/AI, Minor in Management

Aug 2022 — May 2025

Relevant Coursework: Data Structures and Algorithms, Computer Architecture, Intro to AI, Analysis of Algorithms, Operating Systems, Databases, Systems Programming, Data Visualization, Data Mining

WORK EXPERIENCE

Software Engineering Collaborator

Insighter – Startup

April 2025

- Developed a web scraping system in **Python** using **Selenium**, **yt-dlp**, and **BeautifulSoup** to pull meeting videos, agendas, and metadata from city websites, standardizing over **100** meetings into JSON.
- Built reusable scraper modules with **regex** and **NLP-based title matching** to accurately link meeting videos to agendas, improving city onboarding speed and enabling reliable downstream use for **summarization**, **transcription**, and **search features**.

Research Assistant

Purdue University

Sep 2023 – May 2024

- Performed **NLP analysis** on **1,000+ Federal Reserve speeches** using **GPT-based summarization** and **sentiment classification** to predict economic signals.
- Used **OpenAI API**, confidence scoring, and topic modeling (**TF-IDF**) to classify speeches into **hawkish**, **dovish**, or **neutral**.
- Automated policy extraction pipeline with **Python**, **Pandas**, and **OpenAI Batch** processing for faster analysis.

Machine Learning Collaborator

AppUnik – AI Solutions and Predictive Models

Sep 2022 – Aug 2023

- Built and deployed a **GPT-powered recommendation system** using **FastAPI**, **FAISS**, and **Hugging Face Transformers**.
- Implemented hybrid recommendation logic (**collaborative + content-based**) to boost personalization accuracy by **18%**.
- Reduced **API response latency** by **40%** using **Redis caching**, vector search, and model optimization on **AWS Lambda**.

Software Engineering Intern

Battelle – Biodetection Algorithms

Aug 2022 – May 2023

- Improved hazard detection accuracy by **25%** through **Python-based anomaly detection** and signal modeling.
- Reduced false positives by **20%** using iterative validation and cross-checks with ground truth data.
- Delivered modules in **AGILE sprints** and deployed code on **field test platforms**.

PROJECTS

Spotify Clone + AI Music Summary | React, Redux, Express, MongoDB, Node.js, Docker, AWS

- Developed a **full-stack** music streaming app with **user authentication**, **playlist creation**, and playback UI using **React**, **Redux**, and **Firestore Auth**.
- Integrated **GPT-4** to generate AI-based music summaries and playlist captions via **Node.js backend** and **Hugging Face APIs**.
- Containerized the app with **Docker**, deployed using **Amazon Web Services (AWS) EC2**, and automated CI/CD pipelines using **GitHub Actions**.

Fantasy Premier League AI Chatbot | Python, Flask, SQLite, Selenium

- Developed an **AI-powered chatbot** to deliver **FPL player recommendations**, captain picks, and fixture analysis using **Random Forest** and scraped data.
- Created backend **REST APIs** with **Flask** and automated data scraping using **BeautifulSoup** and **Selenium**.
- Optimized data handling with **SQLite** and **Pandas** for **real-time query response** and caching.

Custom Unix Shell | C++, C#, Lex, Yacc, Unix Syscalls

- Developed a **bash-like shell** supporting **piping**, **redirection**, and **background execution**, with parsing via **Lex/Yacc**.
- Implemented features including **quotes**, **escape characters**, **environment variable expansion**, **control structures**, and a line editor with **history**, **tab completion**, and custom prompts.

TECHNICAL SKILLS

Languages: Python, Java, C++, JavaScript, TypeScript, SQL, R, Bash

Frameworks/Libraries: React.js, Node.js, Express, Flask, TensorFlow, PyTorch, Hugging Face, Pandas, NumPy, scikit-learn, Selenium, BeautifulSoup, GraphQL

Tools/Platforms: Git, Docker, Redis, Firebase, AWS (EC2, Lambda, S3), GitHub Actions, MySQL, SQLite, MongoDB, Postman, Jupyter, Linux, Kubernetes

Other: Teaching Assistant for CS193 (Tools), CS252 (Systems Programming), REST APIs, CI/CD, Agile, Authentication Systems