

Hitaansh Jain

New York, NY | (425) 652-7778 | hitaansh1912@gmail.com | github.com/hitaanshjain

Education

New York University

GPA: 3.93/4.0 | BA, Major: Computer Science, Minor: Mathematics

New York, NY

Expected May 2027

- Courses: Software Engineering, Agile Methodology, Data Management and Analysis, Artificial Intelligence, Applied Internet Technology

Professional Experience

Software Engineer Intern

Seattle, WA

Header (joinheader.com)

June 2026 – Present

- Shipped full-stack customization for AI-generated briefings (per-section reorder, toggle, and detail-level control, plus a second-pass LLM-generated TL;DR). Refactored the generation pipeline's transform stage from fixed sections to a prompt assembled dynamically in each user's chosen order and backing it with a JSONB section model and a custom layout editor.
- Owned a release-gating recovery feature end to end, building a paginated, Clerk-authenticated FastAPI endpoint and a frontend screen that let users retrieve past briefing links, and designing it to persist submission metadata through an automated cleanup job so history survived soft-deletion.
- Wrote 130+ automated tests on a test-per-feature workflow while working the full stack, from a multi-provider LLM pipeline and a FastAPI/PostgreSQL backend to a React Native/Expo frontend.

Software Engineer Intern (Part-Time)

Remote

Vardhman Infotech

December 2024 – May 2025

- Engineered backend CRUD logic for a Python inventory management system serving 1,500+ retail locations across 20+ tables, spanning sales, stock, GST compliance, and order tracking with financial-grade data integrity.
- Executed a full SQL-to-Microsoft Access migration, translating a 394-column schema across 19 tables and rewriting 700+ queries across 43 files, resolving type system and schema incompatibilities between backends.
- Refactored a 73-function MySQL database layer into a metadata-driven Access architecture, implementing a configuration-driven schema provisioner that dynamically issued CREATE TABLE and ALTER TABLE operations across 7 .acddb files.

Projects

Local RAG Search Platform

July 2025 – Present

- Built a fully air-gapped RAG system (FastAPI, Llama, ChromaDB) enabling on-device semantic search over private documents, eliminating third-party API data exposure.
- Engineered a hybrid retrieval pipeline fusing vector and keyword search over real legal documents, then diagnosed its ranking flaws through 50-query ground-truth benchmarking across 1,070 pages of zoning ordinances, motivating a BM25 redesign of the keyword scoring.
- Implemented a streaming-response pipeline and utilized the browser Streams API (ReadableStream, TextDecoder) for a real-time token streaming UI, eliminating UI blocking and the wait time for full-generation response payloads.

Stock Analyzer

April 2026 – June 2026

- Designed and deployed a 3-tier stock analysis platform (Docker Compose, Flask, MongoDB) with rate-limit-aware retry logic for OpenAI and SEC APIs, processing jobs across an 8-worker parallel pipeline with 80% test coverage gates enforced via GitHub Actions CI/CD and automated cloud deployment on DigitalOcean.
- Modeled MongoDB data across 6 collections for users, jobs, sessions, screening runs, stock reports, and a global cache; implemented upsert operations ensuring zero data loss across concurrent analysis jobs and user sessions.
- Built session-based authentication and a polling job monitor using Flask, enabling users to launch stock analysis jobs with a live log view updated every 3 seconds.

VocabLearn

February 2026 – May 2026

- Led backend development on a 5-person Agile team, delivering a JWT-authenticated REST API with 13 endpoints and 21 integration tests covering auth, search, quiz, and CRUD flows in Node.js/Express.
- Engineered a multiple-choice quiz engine that samples words from a user's personal word bank, constructs distractor sets from saved definitions, and persists per-word accuracy to MongoDB with dual-layer deduplication (server-side Map + client sessionStorage) to prevent score double-counting.
- Integrated Google Gemini 2.5 Flash for reverse-dictionary lookups, enforcing structured JSON output to enable semantic search-by-meaning over a user's personal word bank.
- Configured a 3-workflow GitHub Actions CI/CD pipeline (backend lint/test with ephemeral MongoDB, frontend lint/build, automated Render deployment), gating all merges to master on 21 passing tests across both services.

Computer Skills

Languages & Frameworks: Python (FastAPI, Flask), Java, JavaScript (Node.js, Express, React), SQL, MongoDB, Pytest, Mocha

AI & Data Engineering: LLMs (OpenAI API, Ollama), RAG (Retrieval-Augmented Generation), LangChain, ChromaDB, Pandas, NumPy

DevOps & Workflow: Claude Code, Git, GitHub Actions, Render, CI/CD, Agile/Scrum