

# ARMAAN AMATYA

📞 (346) 434-0693 ✉️ armaanamatya2014@gmail.com 🔗 [linkedin.com/in/armaan-amatya](https://www.linkedin.com/in/armaan-amatya) 🌐 <https://github.com/armaanamatya>

## Education

### University of Houston

*Bachelor of Science in Computer Science and Mathematics*

**Expected Graduation: December 2026**

*Houston, TX*

## Work Experience

### AlgoVerse

**Aug 2025 – Nov 2025**

#### *Incoming ML Research Intern*

*Remote*

- Working with mentors and peers from Stanford, Berkeley, and Vanderbilt on publishing scholarly research on LLM applications

### AI Agent Stealth Startup

**May 2025 – Present**

#### *Fullstack AI Software Engineer*

*Remote*

- Led the development** of an agentic, AI-powered data analysis platform leveraging a multi-node **LangGraph** architecture
- Built a full-stack web application with a **React 18** and **TypeScript** frontend using shadcn/ui, and an **asynchronous FastAPI** backend with **WebSocket** support for real-time updates, **Motor** for non-blocking MongoDB access, and **MongoDB** for storage and session persistence
- Engineered dynamic LLM context windows and custom prompt flows, improving structured output accuracy and reducing hallucinations by **50%**, and increasing multi-step task success by **30%**
- Designed a scalable session persistence system with custom LangGraph checkpointing, TTL-based resource cleanup, and an optimized MongoDB schema, cutting state recovery time from **5s to under 1s**
- Implemented a secure file upload and storage system using **Google Cloud Storage**, with support for structured data ingestion via **BigQuery**
- Containerized and deployed the application using **Docker** and **Google Cloud Platform**

### MyMCAT.ai

**Aug 2024 – April 2025**

#### *Software Engineering Lead*

*Houston, TX*

- Led the development** of a AI-based education platform for the MCAT, serving **1,000+ university students**, utilizing **NextJS** and **TypeScript** for the frontend and backend, **Prisma** as the ORM, **Supabase** for the database, **Clerk** for User Authentication and **Stripe** for payment/subscription handling

### FuseMachines

**May 2024 – Aug 2024**

#### *Machine Learning Engineer Intern*

*New York, NY*

- Engineered **Large Language Models (LLMs)**, including **BERT** and **Attention Models**, to analyze software engineering requirements and generate accurate dependency lists, improving project planning efficiency by **37%**
- Analyzed **1200+** software requirements monthly using **scikit-learn (classification models)**, **spaCy**, and **NLTK** in **Python**, saving the company **\$2000 annually** by decreasing project lead times
- Integrated **TF-IDF vectorization** to detect similarities in software requirements, focusing on **preprocessing, model training, hyperparameter tuning** and **clustering** to optimize **NLP** analysis processes

### F1Soft International

**May 2023 – Aug 2023**

#### *Software Engineer Intern*

*Remote*

- Engineered high-performance **RESTful APIs** using **Java Spring Boot**, slashing request latency by **50%**
- Implemented a secure and stateless user authentication system in **Java Spring Boot**, leveraging **Spring Security** and **JWT** for login, with user signup, email verification, authenticated user retrieval, and robust error handling; integrated with **Supabase** and validated endpoints using Postman

## Technical Skills

**Languages:** Python, C++, Java, JavaScript, TypeScript, SQL

**Developer Tools:** Google Cloud Platform, Postman, PowerBI, Git, Docker, Jira

**Libraries/Frameworks:** ReactJS, NextJS, NodeJS, ExpressJS, Spring Boot, FastAPI, Streamlit, Flask

**Databases:** MySQL, PostgreSQL, MongoDB, Supabase, Redis, VectorDB

**AI/ML:** Pytorch, Keras & TensorFlow, LangChain, LangGraph, vLLM, HF Transformers, Scikit-learn, Matplotlib, Numpy, Pandas