Hsiang Yu Huang (Anna)

Allston, MA | Tel: +1 617-319-5044 | huanganna1004@gmail.com | www.hsiangyuhuang.com | linkedin.com/in/hsiangyuhuang

EDUCATION

Boston University   Boston, MA
Master of Science in Data Science | GPA: 3.59 / 4.0   *Dec. 2025*
Relevant Courses: Deep Learning, Graduate Databases, Data Engineering, Time Series, Artificial Intelligence, Crypto

National Taiwan University of Science and Technology   Taipei, Taiwan
BBA in Industrial Management and Bachelor Program of Finance, Minor in Computer Science | GPA: 3.85 / 4.3   *Jun. 2023*
Relevant Courses: Algorithms, Machine Learning, Data Analytics, Statistics, Object-Oriented Programming

SKILLS

- **Languages & Frameworks:** Python, SQL, C++, TypeScript, JavaScript, R, FastAPI, Next.js, Node.js, Flask, React, HTML/CSS
- **Cloud & DevOps:** AWS (Lambda, S3, EC2), Azure, PostgreSQL, MongoDB, Docker, Linux, Nginx, CI/CD, Git
- **Data Engineering & Tools:** PySpark, ETL Pipeline Automation, SQL Database Design, Azure Synapse, PostHog, Power BI
- **AI Engineering & ML:** LangGraph, LangChain, MCP, PyTorch, TensorFlow, Scikit-Learn, Pandas, NumPy, NLP, Time Series

WORK EXPERIENCE

Full Stack Developer – Citale (citaleco.com)   Boston, MA
BU Spark! Launch Lab (Stipend) | Stack: Next.js, SQL, Supabase, Google Maps API, Vercel, PostHog   *Sep. 2024 – Present*

- Designed **PostgreSQL schemas** with foreign keys for users, posts, and relationships, ensuring referential integrity.
- Implemented **10+ features** including messaging, notifications, and user profiles using **Supabase** and **Next.js**.
- Deployed production app on **Vercel** with **PostHog** integration for user analytics and system monitoring.

Research Assistant – LLM Platform   Boston, MA
BU BIT Lab (Unpaid) | Stack: FastAPI, Next.js, MongoDB, LangGraph, Prompt Engineering, OpenRouter, MS Clarity   *Sep. 2025 – Present*

- Architected a **RESTful API using FastAPI** to orchestrate a **LangGraph multi-agent system**, managing asynchronous communication between Coordinator, Writer, and Reviewer agents.
- Designed a **MongoDB** schema to store complex conversation history and retrieval contexts, ensuring efficient data persistence for the multi-model console.

Research Assistant – Machine Learning for Sales Forecast   Taipei, Taiwan
NTUST Artificial Intelligence and Decision Analysis Lab | Stack: R, XGBoost, Time Series Analysis   *Nov. 2023 – Jul. 2024*

- Built an **ARIMA–XGBoost** forecasting pipeline, boosting predictive accuracy (R^2) from **8.3% to 73.4%**.
- Developed a **conditional rolling window model** to enable adaptive, real-time predictions.

PROJECTS

More Projects are available here: hsiangyuhuang.com/projects

Boston Weekend Agent   Boston, MA
Stack: AWS (Lambda, Step Functions, EC2, S3), Python, LangChain, Cron   *Nov. 2025*

- Orchestrated a serverless ETL pipeline using **AWS Step Functions** and **Lambda** to ingest real-time event data.
- Deployed an **EC2**-based web scraper with **Cron** jobs to bypass anti-bot restrictions for reliable data extraction.
- Generated AI summaries via **LangChain** and stored static JSON in **S3** for low-latency frontend delivery.

From Tweets to Trends – Predicting Stock Volumes Using X Sentiment   Boston, MA
Course: Big Data Engineering | Stack: Azure Functions (Python), ADLS Gen2, Rapid API   *Feb. 2025 – May. 2025*

- Leveraged **Azure Functions (Timer Trigger)** to establish **Cron-based automation** for the daily ingestion of social media streams from X (Twitter).
- Executed the **Extract & Transform (ETL)** process using **Python** to parse raw JSON data, converting unstructured logs into analysis-ready formats.

Winner – DS+X Hackathon 2025 (Best Overall, HackBU 1st)   Boston, MA
BU Spark! | Stack: Python, FastAPI, OpenAI API, OpenAlex API   *Oct. 2025*

- Engineered the backend for **RhettSearch**, an interactive research gamification platform driven by AI.
- Orchestrated a **RAG pipeline** that retrieves literature via **OpenAlex API** and synthesizes insights using **OpenAI API**.
- Designed structured **RESTful APIs** to parse complex JSON citation data, ensuring serialization for frontend integration.