Hsiang Yu Huang (Anna)

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

SUMMARY

Data Science graduate student at Boston University with hands-on experience in data preprocessing, exploratory analysis, and predictive modeling. Proficient in Python, SQL, AWS, and Azure pipelines for data engineering and machine learning workflows. Skilled in ETL automation, feature engineering, and statistical analysis to deliver data-driven insights and scalable solutions. A collaborative problem-solver dedicated to transforming complex datasets into actionable intelligence that supports decision-making. **Seeking Data Science, Data Engineer, Data Analytics, or Software Engineer Full-Time Job (Jan 2025).**

EDUCATION

Boston University Boston, MA

Master of Science in Data Science | GPA: 3.57 / 4.0 | Expected Dec. 2025

Relevant Courses: Deep Learning, Graduate Databases, Data Engineering, Time Series

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

WORK EXPERIENCE

Full Stack Developer – Citale Boston, MA

BU Spark! Launch Lab *Jan. 2025 – May. 2025*

- Selected for BU Spark! Launch Lab's Volunteer Track, receiving a $500 stipend for prototype development.
- Built a Boston-based social platform enabling event discovery and local networking.
- Developed frontend in React and backend logic in SQL, integrating Google Maps API for event visualization.
- Collaborated with a cross-functional team to achieve milestone deliverables and usability improvements.

Research Assistant – Machine Learning for Sales Forecast in Graphic Card Manufacturing Taipei, Taiwan

NTUST Artificial Intelligence and Decision Analysis Lab *Nov. 2023 – Jul. 2024*

- Built an ARIMA–XGBoost forecasting pipeline, boosting R^2 from 8.3% to 73.4%.
- Developed a conditional rolling window model for adaptive, real-time predictions.
- Supported procurement and inventory planning using data-driven sales insights.

Research Assistant – Smart Vending Machine Shelf Optimization Taipei, Taiwan

NTUST Decision Analysis and Applied Statistics Lab *Apr. 2023 – Sep. 2023*

- Clustered product sales with K-means using metrics like mean, CV, revenue, and unit price.
- Built a classification tree to identify product-shelf performance patterns by price segment.
- Delivered actionable recommendations to improve shelf planning and profit optimization.

PROJECTS

Efficient Open-Vocabulary Models for Low-Power Computer Vision (LPCV Competition) Boston, MA

Course: Deep Learning *Feb. 2025 – May. 2025*

- Optimized X-Decoder using DyT, SwiGLU, and linear attention to reduce inference cost.
- Evaluated on COCO and RefCOCOg datasets to align segmentation performance with low-power device requirements.
- Achieving a 7.5% GPU usage reduction and improving segmentation accuracy from 17 to 22 mIoU.

From Tweets to Trends – Predicting Stock Volumes Using X Sentiment Boston, MA

Course: Big Data Engineering *Feb. 2025 – May. 2025*

- Built a real-time pipeline to forecast NVIDIA's stock volume using sentiment data from tweets.
- Leveraged Azure Synapse, RapidAPI, and Power BI in a Medallion Architecture framework.
- Connected social media signals with trading behavior for adaptive forecasting and risk analysis.

Equity in Federal Budget Earmarking Processes Boston, MA

Course: Tools for Data Science *Sep. 2024 – Dec. 2024*

- Extracted and cleaned earmark funding data from Senate PDFs using Python (Camelot).
- Analyzed demographic equity in federal funding allocations and visualized funding disparities across Massachusetts.

LLM Platform – Multi-Agent Orchestration & Retrieval System Boston, MA
BIT Lab *Sep. 2025 – Now*

- Built a FastAPI-based LLM logging + retrieval backend capturing prompts, responses, embeddings, intent classes, and provider metadata to support semantic search and interaction analytics.
- Developed a React/Next.js LLM console UI with provider-based state, model selector, and real-time visualization of prompt/response flows across multi-model orchestration (OpenAI, Anthropic, OpenRouter).
- Designed a product-enrichment workflow where LLM outputs are parsed for product mentions, matched with real product data via Google SERP API, and surfaced as interactive product cards on the website.

Modeling Power System Frequency Regulation Reserve Trading Volume Taipei, Taiwan
NTUST Decision Analysis Lab *Sep. 2021 – May. 2022*

- Forecasted electricity demand using SARIMAX and Backpropagation Neural Networks (BPN), achieving 69.4% accuracy.
- Identified time-based and holiday-driven consumption patterns for grid optimization.

ADDITIONAL EXPERIENCE

Research Volunteer – Boston University BIT Lab (Business Insights through Text) Boston, MA
Sep. 2025 – Present

- Participating in ongoing research on Generative Engine Optimization and Marketing Science under the supervision of Prof. Dokyun Lee.
- Conducting literature reviews on generative engine optimization and its industry applications.
- Preparing to contribute to forthcoming empirical studies on generative AI and marketing analytics.

Creator – Boston Weekend Agent (Personal Project) Boston, MA
Oct. 2025 – Present

- Designed and deployed a serverless **AWS Step Functions** pipeline integrating **Lambda**, **S3**, and **EC2** for automated Boston weekend event summaries.
- Automated data retrieval, LLM-based text generation, and cloud report storage for fully managed content delivery.
- Implemented scalable workflow orchestration and cloud automation for reliable data-driven report generation.
- View live reports at **hsiangyuhuang-anna.vercel.app/weekend_report**.

Winner – DS+X Hackathon 2025 (Best Overall, HackBU 1st) – BU Spark! Boston, MA
Oct. 2025

- Developed **RhettSearch**, an interactive research gamification platform connecting BU students with AI-driven paper discovery.
- Integrated semantic search, user gamification, and AI-generated recommendations using OpenAI API and OpenAlex API.

Winner – Civic Tech Hackathon 2025 (Hack the Herbaria) – BU Spark! Boston, MA
Feb. 2025

- Created **U.S. Virtual Garden**, a website with an interactive dashboard visualizing national herbaria data, integrating the Groq API for dynamic species descriptions.
- Implemented user profiles and a "random box" gamification feature, leveraging LLM APIs to generate personalized plant descriptions

SKILLS

- **Programming & Tools:** Python, SQL, R, C++, JavaScript, React, Git, Linux, HTML/CSS, ETL, Data Pipeline Automation
- **Machine Learning & Analytics:** PyTorch, TensorFlow, Scikit-Learn, Pandas, NumPy, Feature Engineering, Predictive Modeling, Time Series Analysis, NLP, Statistical Analysis
- **Database & Cloud:** Azure, AWS(Lambda, S3, EC2, Step Functions), Supabase, SQL Database Design, Query Optimization
- **Visualization & Communication:** Power BI, Looker Studio, Matplotlib, Seaborn, Data Storytelling, Dashboard Design