

Huan Shuo Hsu

Email: huan246shuo@gmail.com | hh3038@columbia.edu Tel: (347)-727-8391

LinkedIn: <https://www.linkedin.com/in/huan-shuo-hsu>

EDUCATION

Columbia University - *M.S. in Applied Analytics*

Sep. 2024 – Dec. 2025

- **Courses:** Machine Learning, Data Modeling, Predictive Analysis, Cloud Computing

Pratt Institute - *B.Arch. in Architecture | Magna Cum Laude, Dean's List*

Aug. 2018 – May 2023

- **Courses:** Statistics, Calculus, Physics, Structural Mechanics, Architectural Design, Creative Coding

SKILLS

- Programming Languages: Python, R, C, HTML/CSS
- Tools: SQL, NoSQL, Tableau, Power BI, Docker, Tensorflow, Pytorch, Pandas, Matplotlib, Apache Kafka, AWS EC2, S3, Glue, Athena
- Languages: English (Native), Mandarin (Native)

WORK EXPERIENCE

Pulse Protocol

Pasadena, CA

UX Researcher / Data Analyst

May. 2022 – Sep. 2022

- Designed and deployed a Medical Tourism website into functional, responsive web pages using HTML and CSS.
- **Improved conversion rate for medical procedures by 10%** using Tableau for exploratory data analysis.
- **Boosted the usage of critical app features by 20%** by conducting analysis using Google Analytics and SQL.

Stardust Lab

Remote

Project Lead / Linux Administrator

Feb. 2020 – May 2022

- **Lead a team** of software engineers and designers to wireframe, develop, lead sprints, and weekly stand-ups.
- **Maintained over 2,000 active players** throughout the project duration and **generated over \$50,000 in revenue**.
- The project was later acquired by another development team.

Atelier-3

Taipei, Taiwan

Data Analyst

May. 2019 – Sep. 2019

- Measured deviation from previous project timelines using Excel, analyzing historical data to **improve time efficiency by 30 days**, and **achieved a 10% reduction in project duration**.

PROJECT EXPERIENCE

Click Through Rate Prediction with XGBoost

Data Science & Machine Learning

Oct. 2024 – Nov. 2024

- Leveraged **XGBoost with Bayesian Optimization** for **hyperparameter tuning**, achieving a significant reduction in RMSE for Click-Through Rate prediction.
- Leveraged statistical tests and importance scores from **linear regression, decision tree, random forest** to systematically refine **feature engineering**, improving model accuracy and reducing computational complexity.

Live Query of Stock Market Price Change with Kafka

Data Engineering

July 2024 – Sep. 2024

- **Developed a real-time data pipeline** to stream stock market data using Python and Apache Kafka.
- **Deployed Kafka on EC2 instances**, processed and **stored stock market data in AWS S3**, and implemented process **optimization utilizing AWS Glue** to catalog the data for efficient **querying with AWS Athena**.
- **Automated the real-time ingestion and processing pipeline with Zookeeper**, ensuring continuous updates and high system reliability.