

KERSTIN (YU HAN) SUN

Houston, TX 77030 | 510-996-6090 | kerstinsun2000@gmail.com | [linkedin.com/in/yuhanssun](https://www.linkedin.com/in/yuhanssun) | github.com/yuhansun33

Rice University Master of Computer Science with experience in Data Engineering and AI/ML.

Track record in architecting end-to-end data solutions and advanced models at Amazon.

EDUCATION

Rice University - Top 20 University in the U.S.

Houston, TX

Master of Computer Science

Aug 2025 - Dec 2026 (Expected)

National Yang Ming Chiao Tung University - Top 3 University in Taiwan

Hsinchu, Taiwan

Bachelor of Science in Industrial Engineering and Management

Sep 2020 - Jun 2024

- GPA: 3.97/4.0; Ranking: 2/72; Academic Achievement Award
- Coursework: Data Structures and Algorithms, Database Systems, Artificial Intelligence Capstone, Operating Systems

WORK EXPERIENCE

Amazon - Scaling a \$1B+ Seller Ecosystem with Data Science and Automated Data Pipelines

Taipei, Taiwan

Data Analyst

Jul 2024 - Jul 2025

- **Machine Learning:** Drove a **60%** acceleration in new seller launch time by engineering a segmentation model using K-Means **clustering** and validated distinct seller profiles via ANOVA to prioritize high-potential targets
- **Data Pipeline Architecture:** Designed an automated ETL pipeline using **Python** and **SQLite** to transform raw data into structured datasets, powering weekly business reviews and reducing manual reporting time for stakeholders by **40%**
- **Data-Driven Strategy:** Led 5 teams to launch a **Go-to-Market** (GTM) strategy by analyzing customer behavior and sales data, leading a Taiwan Select campaign that drove **3,000+** unique clicks for 100+ sellers during Prime Day
- **Marketing Analytics Solution:** Increased ad spend by **13%** by developing a marketing analytics framework and delivering **30+** data-driven consultations based on campaign performance and analysis

Syntec Technology - Architecting Data Solutions for a Leader in Industrial Automation

Hsinchu, Taiwan

Data Engineer Intern

Jun 2023 - Nov 2023

- **Enterprise Data Integration:** Built a data pipeline to integrate data from SAP system into a centralized dashboard that visualized production and material data to enhance cross-departmental operational visibility for **50+** stakeholders
- **Intelligent Automation:** Developed a Robotic Process Automation (RPA) solution to automate manufacturing workflows, which reduced routine work hours by **7%** and ensured higher data integrity
- **Data Analytics Scripting:** Developed a VBA **automation script** to create a packaging assignment tool, which analyzed item volume and weight data to automate tasks and increase production line operational efficiency by **10%**

PROJECTS

Comparative Machine Unlearning Analysis on CNNs | PyTorch, VGG-16, ResNet-18, SalUn

Sep 2025 - Dec 2025

- Evaluated **machine unlearning** efficacy across VGG-16 and ResNet-18 by implementing Gradient Ascent and SalUn on CIFAR-10, establishing empirical benchmarks for model selection in **GDPR** and **privacy-sensitive** applications

Intelligent Tutoring System with Multi-Agent LLM | CrewAI, LangChain, Vector DB

Feb 2024 - Jun 2024

- Developed a **multi-agent** tutoring system with CrewAI, integrating **RAG** and **Chroma DB** for evidence-based answers
- Benchmarked LLaMA-3 and GPT-4o, and optimized the retrieval system with GPT-4o and OpenAI embeddings to improve semantic stability and cut redundant API calls by 20%

Full-Stack Blockchain Supply Chain Application | React, Web3.js, Express, MySQL

Feb 2023 - Mar 2024

- Engineered a **full-stack** supply chain tracking platform, developing a React frontend for real-time interaction with smart contracts via Web3.js, and an Express.js backend with MySQL for secure user authentication
- Deployed smart contracts on a simulated **blockchain** network using Truffle and Ganache

TECHNICAL SKILLS

General Skills: Data Engineering, Data Science, Machine Learning, Deep Learning, Optimization

Languages: Python, SQL, C/C++, C#, HTML, CSS, JavaScript

Tools & Frameworks: Hadoop, Spark, AWS, GCP, PyTorch, TensorFlow, SQLite, Pandas, React, Express, Git, GitHub