

Riverside, CA 92507 | (323)348-3616 | hyao014@ucr.edu | <https://www.linkedin.com/in/daniel-yao182124/>

Portfolio Website: <https://portfolio-website-rho-three-40.vercel.app/>



Objective:

Data Science graduate with hands-on experience in machine learning, NLP, and end-to-end project development. Proven ability to build predictive models, automate workflows, and translate complex datasets into actionable business insights. Experience includes leading sustainability-focused data initiatives and developing scalable backend systems. Bilingual in Mandarin and English with a strong foundation in Python, SQL, and statistical modeling.



Education:

[University of California, Riverside](#)

Fall 2022 – March 2025

Major: Bachelor of Science in [Data Science](#)



Work Experience:

June 17th – August 9th 2024

➤ Data Science Fellow - UN SDG Automation Project (*NLP, Regex, Semantic Search*)

- Led the development of a course classification system used to align over **10,000 university courses** with the **UN Sustainable Development Goals (SDGs)**, directly supporting **institutional funding eligibility** through STARS reporting.
- Engineered high-accuracy pipelines using **Python, regex, Elasticsearch**, and **large language model (LLM)** embeddings, achieving **80%+ classification precision** and reducing manual reporting efforts by 70%.
- Delivered consistent **95%+ data accuracy** through scalable data extraction and transformation workflows sourced from registrar and public-facing course portals.
- Established a sustainable, **automated reporting framework** that empowered universities to track progress on SDG-related initiatives with greater consistency and transparency



Projects:

➤ Project Leader- Text-Based RPG (*C++*)

- Directed a 4-person development team to design and deliver a fully functional RPG game, completing the project 12.5% ahead of schedule.
- Spearheaded gameplay concept design, delegated coding responsibilities, and ensured seamless integration and testing of core features.
- Achieved a 98% evaluation score, reflecting the project's technical quality, user engagement, and team coordination.

➤ Software developer - E-Commerce Ordering System (*Python/SQL*)

- Engineered a full-featured backend system for an Amazon-like ordering platform, enabling user registration, order tracking, and multi-store purchasing.
- Integrated a **SQL**-based data storage layer to optimize data access efficiency, improving system performance by 25% and reducing potential server load by 15%.
- Collaborated on **front-end** flow and **UX structure**, reducing navigation time by 30% and shortening order processing time by 20%, which could contribute to a 10% boost in customer satisfaction and a 5% increase in potential sales.

➤ Bank Marketing Data Analysis (*R, Python, multiple Regression models*)

- Built a **logistic regression model** using **one-hot encoded** marketing campaign data to predict term deposit subscriptions, **revealing key conversion factors** such as call duration and previous outcome.
- Delivered actionable insights that informed strategy for customer segmentation and **outreach optimization**, demonstrating the business impact of **predictive modeling**.



Relevant Skills:

- **Languages:** Python, R, SQL, C++, Java, PostgreSQL, Git
- **Libraries & Tools:** Pandas, NumPy, Matplotlib, Scikit-learn, Regex, Elasticsearch
- **Machine Learning & Analysis:** Logistic Regression, Clustering, Data Cleaning, Predictive Modeling, Semantic Search (LLM Embeddings)
- **Other:** Git, Excel, Bilingual – English and Mandarin Chinese