

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



### Objective:

Detail-oriented Data Science student with strong experience in **Python, machine learning**, and sustainability-related **large dataset** projects. Passionate about **organizing information, solving problems**, and **delivering practical solutions**.



### Education:

[University of California, Riverside](#)

Fall 2022 – March 2025

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



### Work Experience:

- **Data Science Fellow - Sustainability Project** June 17<sup>th</sup> – August 9<sup>th</sup> 2024
  - **Web Scraping:** Extracted and organized course data from university portals (e.g., Rweb, Coursicle) and the Registrar, processing over 10,000 course records with an accuracy rate of 95%.
  - **Method Development:** Implemented text pattern matching using Python **regex** and **Elasticsearch** to identify relevant courses, increasing classification efficiency by 60%. Additionally, utilized **LLMs** to embed SDG keywords, definitions, and course descriptions, and applied cosine similarity, achieving an over 80% match accuracy when comparing course content to SDG definitions.
  - **Outcome:** Streamlined the course-SDG matching process, providing universities with tools to generate consistent, accurate, and comprehensive SDG reports. The system potentially improves report generation speed by 70%, supporting global sustainability initiatives while ensuring data consistency and accuracy.



### Projects:

- **Project Leader- Text-Based RPG (C++)**
  - Led a 4-person project group in developing a text-based role-playing game, completed the project 12.5% faster than expected.
  - Spearheaded the conceptualization and initial idea development. Responsible for **assigning coding work, integrating codes, debug, and testing**.
  - Achieved 98% in project evaluation, underscoring its success and seamless execution.
- **Software developer - E-Commerce Ordering System Clone (Python/SQL)**
  - Collaborated closely with another team member to swiftly develop the system, focusing on both functionality and user interface optimization, improving user navigation speed by 30%.
  - Utilized **SQL database** technologies to ensure efficient data storage and retrieval, enhancing system performance by 25%, which could potentially reduce server costs by 15% if scaled for Amazon-level operations.
  - The optimized workflow could reduce order processing time by up to 20%, potentially leading to a 10% increase in customer satisfaction and a 5% increase in overall sales for similar e-commerce platforms.
- **Bank Marketing Data Analysis**
  - **Analyzed a large dataset** related to direct marketing campaigns (phone calls) of a Portuguese banking institution.
  - Conducted **data cleaning, one-hot encoding, and logistic regression** modeling to predict client subscription to a term deposit. Collaborated with a group to develop and implement predictive models based on the dataset, contributing to insights for marketing strategy optimization.



### Relevant Skills:

- **Technical Skills:**
  - **Python, R, C++, Java, Matplotlib**, and **AI/ML/DM** for data analysis, data visualization. **MySQL, PostgreSQL**. Experienced in **Excel, Pandas**, and **NumPy**. **Marketing Research/Market Research, Marketing Strategy**
- **Language Skills:**
  - **Bilingual: Mandarin, English.**