

HANG LYU

San Diego, CA

(858) 531-9406

www.linkedin.com/in/lyuhang

halyu@ucsd.edu

EDUCATION

Master of Science in Business Analytics, Rady School of Management

12/2025

University of California, San Diego, CA

- Relevant Courses: AI-Math and Programming for Business Analytics, SQL and ETL, Collecting and Analyzing Large Data, Web Mining and Recommender System, AI-Assisted Customer Analytics.

Bachelor of Science, Applied Mathematics

06/2024

University of California, San Diego, CA

- Relevant Courses: Probability, Intro to Python, Applied Linear Algebra, Intro to Java, Vector Calculus, Mathematical Reasoning, Numerical Methods/Physical Modeling, Elements of Complex Analysis.

EXPERIENCE

Data Analyst Intern Leader, Capgemini Co. Ltd, Shanghai, China

07/2023 - 09/2023

A global consulting, technology services, and digital transformation company that operates in IT and consulting industry.

- Led a 20-person data team on an EV sentiment analysis project, building a scalable end-to-end pipeline from data collection to BI output.
- Developed a multithreaded crawler in Python and Selenium to extract 700K+ records (e.g., reviews, complaints) from 14 major auto platforms, including(e.g., Autohome, Sohu Auto).
- Architected a dual-database system (MySQL + MongoDB), accomplishing field normalization and boosting processing efficiency by 2x.
- Co-developed crawler and visualization features to support sentiment analysis delivery, collaborating with planning and algorithm teams across key technical milestones.

PROJECTS

Feature-Based Similarity Analysis for Amazon Health Products, Web Mining and

09/2024 - 12/2024

Recommender System, UCSD Rady School of Management

- Conducted exploratory analysis of 494,121 reviews and 62,597 products using Python (Pandas), uncovering key consumer behavior insights during pandemic.
- Designed a Jaccard Similarity-based recommendation system to identify most similar products for target item, achieving a max similarity score of 1.0 after feature optimization.
- Evaluated collaborative filtering and keyword models on 460K+ users; visualized user preferences to facilitate recommender tuning.

SQL Database Project, SQL and ETL, UCSD Rady School of Management

08/2024 - 09/2024

- Managed a 4-member team to structure 42 CSV and XML files containing purchase orders and supplier data into Snowflake.
- Built a scalable ETL pipeline using Python and Snowflake to process 9K+ supplier transactions, resolving field and timestamp inconsistencies.
- Integrated 105K+ ZIP codes from Snowflake Marketplace to support supplier performance analysis across regions.

SPECIALIZED SKILLS

- Languages/Coding: Python (Pandas, NumPy), SQL (Advanced), Java, JavaScript, R, MATLAB, Spark.
- Machine Learning & AI: TensorFlow, Keras, Scikit-learn, Linear/Logistic Regression, Neural Networks, NLP.
- Data Processing & Analytics: Data Wrangling, Feature Engineering, Statistical Modeling, Predictive Analytics.
- Databases & Big Data: MySQL, SQLite, MongoDB, Snowflake, Hadoop (HDFS).
- Visualization & Tools: Power BI, Tableau, Matplotlib, Seaborn, SQL Developer, Docker, AWS, ETL, A/B Testing.

PROFESSIONAL AFFILIATIONS & LEADERSHIP

- Teamed Member: Rady Consulting Club.

08/2024 - Present

- Elected Chair: Celebrating Muir College Week; Muir College Council.

11/2022 - 06/2023

Hosted a 300+ event and leveraged Excel to calculate \$4808.72 budget.