

# YIHAN ZHOU

San Francisco, CA ◊ (+1)510-679-9036 ◊ [yihan.zhou427@gmail.com](mailto:yihan.zhou427@gmail.com) ◊ [LinkedIn](#) ◊ [GitHub](#) ◊ [Profile](#)

## PROFESSIONAL SUMMARY

**Applied Data Scientist and AI Engineer** specializing in building end-to-end analytics and ML-driven decision-support systems. Experienced in translating ambiguous business challenges into scalable data products deployed in production, delivering measurable impact across consulting, industrial R&D, and quantitative finance.

## AREAS OF EXPERTISE

Applied Machine Learning ◊ Statistical Modeling ◊ Decision-Support Systems ◊ Analytics Product Development  
End-to-End Data Pipelines ◊ Production ML ◊ Client-Facing Analytics ◊ Executive Communication

## PROFESSIONAL EXPERIENCES

### Data Scientist

*McKinsey & Company*

Sep 2025 - Present

San Francisco, CA, U.S

- **Automated Workforce Planning:** Architected “SWP Lite”, an end-to-end platform integrating industry benchmark data, a *SQL analytics agent*, and *Python* pipelines, enabling self-serve scenario modeling and analytics for enterprise clients; generated **\$140K in client project revenue**
- **AI-Driven Taxonomy Mapper:** Engineered an *LLM-powered* system to standardize fragmented job role data, achieving **> 90% accuracy** and **> 0.70 F1**; system monetized via a **\$15K internal study** and adopted by **11 enterprise clients**
- **Operational Talent Analytics:** Conducted end-to-end data pipelines across HRIS, recruiting, and absenteeism data, delivering a reusable *PowerBI* dashboard asset to define operational KPIs and surface productivity insights under ambiguous business requirements

### Data Scientist, Research & Development

*Gotion Inc.*

Jun 2025 - Aug 2025

Fremont, CA, U.S

- **Predictive ML Pipelines:** Built production-ready pipelines and feature engineering workflows to predict battery grading capacity and resistance at early manufacturing stages, enabling **immediate grading evaluations**
- **Multi-stage XGBoost Modeling:** Developed a hierarchical modeling approach combining classification-based segmentation, segmented regression, and post-prediction calibration, achieving 100% of predictions within 1% absolute relative error and 92.6% within 0.5%, with results contributing to **2 submitted patents** and **2 peer-reviewed publications**
- **Cost Optimization:** Delivered significant operational impact, saving **\$500K per GWh**, reducing grading time from **6 hours to 0**, and increasing battery throughput by **32%**

### Data Engineering Intern (Large Model Division)

*SenseTime, IT Services and IT Consulting*

May 2024 - Aug 2024

Beijing, China

- **Generative AI System:** Designed a text-to-poster generation pipeline converting user prompts into production-quality visual assets; **reduced operational costs by 78%** and forecast **300K+ user growth**
- **Automated Data Harvesting:** Built a *Python Selenium*-based pipeline to collect 1,000+ PDF training assets, **lifting LLM model precision by 13%**
- **LLM Benchmarking:** Implemented an automated evaluation framework via *GPT APIs* and persona testing; boosted satisfaction from **20%** to **70%** through strategic model fine-tuning

### Data Engineering Intern (Data Group, Excellent Intern)

*ZhuoShi Fund, Financial Services Hedge Fund*

Feb 2024 - May 2024

Beijing, China

- **Trading Data Infrastructure:** Built *production-grade* data pipelines for quantitative trading systems, standardizing **24 years** of Hong Kong Exchange data and detecting **175+ data anomalies** to ensure reliable downstream decision-making
- **Multi-Exchange Scraper:** Automated daily settlement parameter scraping from **6 exchanges** by creating a scalable data extraction pipeline, allowing flexible querying by exchange, data, and time period
- **Data Governance Automation:** Streamlined data governance and **saved 7+ hours weekly** for the trading team through automating metadata management with *DataHub*, *Git* and *Airflow*, enhancing data lineage visualization

## EDUCATION

### Master of Analytics - University of California, Berkeley

Industrial Engineering and Operations Research

Sep 2024 - Aug 2025

GPA: 4.00/4.00 (Top 1%)

**Awards:** Excellence in Student Leadership Award

### Bachelor of Science & Engineering - Tsinghua University

Mathematics and Physics, Civil Engineering and Systems

Major GPA: 3.88/4.00 (Top 20%)

Sep 2020 - Jun 2024

**Awards:** Outstanding Graduate, Outstanding Students' Leader, Excellent A-class Comprehensive Scholarship