

# Changming Liu

Boston, MA | 858-648-2594 | changming\_liu@g.harvard.edu

## EDUCATION

### Harvard University

Boston, MA

Master of Data Science

September 2025 – March 2027

### University of California, San Diego (UCSD)

San Diego, CA

Bachelor of Mathematics and Computer Science; Double Major in Business Economics September 2021 – June 2025

**Cumulative GPA:** major GPA 4.0/4.0 (Total GPA: 3.89/4.0) with Provost Honor every semester; GRE: 335/340

**Relevant Courses:** Advanced Data Structures and Algorithms, Software Engineering, Distributed Systems, Web Mining, Artificial Intelligence, Computational Statistics, Recommender Systems, Data Analysis, Theory of Computability

## WORK EXPERIENCE

### CNPC USA

Houston, TX

Software Engineering Intern

July 2024 – October 2024

- Developed a scalable data ingestion pipeline for large geological datasets, leveraging SQLAlchemy with MySQL for efficient structured data storage and retrieval, reducing data preprocessing time by 40%.
- Developed a microservices based real time geological data ingestion system using FastAPI and Node.js, integrating Redis streams to process high frequency sensor and drilling data, reducing pipeline latency by 35%.
- Designed and implemented scalable algorithms using Dask & Pandas, structuring large oil and gas datasets using Python and SQLAlchemy, reducing data preprocessing time and improving downstream ML model accuracy.

### Boston Consulting Group (BCG)

Shanghai, China

Biotechnology Consulting Intern

June 2023 – August 2023

- Developed an optimized bioinformatics data pipeline for genomic analysis, improving data processing speed by 30% and enabling the handling of large scale genomic datasets and deeper genetic variation insights.
- Analyzed over 50 policy documents and conducted 15 expert interviews to evaluate the impact of recent Chinese healthcare reforms on the biotechnology industry. Compiled findings into a strategic report that outlined key risk factors, providing insights for biotech firms to refine their market entry strategies and enhance compliance protocols.
- Collaborated with BCG's analytics team to deploy cloud based machine learning workflows (AWS Lambda, S3) for biotech data modeling, enhancing CRISPR market simulation accuracy across 30+ datasets.

### CITIC Securities

Beijing, China

Data Engineering & API Development Intern

June 2022 – August 2022

- Leveraged regression analysis and Monte Carlo simulations and employed PowerBI for extensive historical data visualization to analyze each fund's risk and return profile which led to a 20% increase in investor inquiries.
- Developed an automated financial news aggregation system using Selenium and FastAPI, implementing data cleaning and structured parsing to extract relevant market insights before storing them using MySQL database.
- Assisted in maintaining and adding features to the company's online platform using Angular, enhancing data validation and integrating real time API communication with FastAPI and Redis caching.

## RESEARCH, PROJECT & LEADERSHIP EXPERIENCE

### Biological Data Analysis Research, UCSD Ion Channel Neurotransmission Analysis Assistant August 2024 – Present

- Developed a Python pipeline using NumPy, Pandas, and OpenCV for microscopy image integration, reducing data processing time by 10 minutes per dataset of 10,000 images, enabling reliable and efficient batch analysis.
- Utilized MATLAB functions like polyfit and findpeaks for polynomial regression and peak intensity analysis, achieving 95% accuracy in calcium intensity peak prediction and extracting precise biological metrics such as half life.
- Developed a web based interactive dashboard using React.js and FastAPI, integrating a RESTful API with MongoDB for structured storage, improving data accessibility and reducing manual report generation time by 40%.

### Full stacked AI based Tarot card reading Web App, UCSD Project Leader

January 2023 – June 2023

- Developed AI based Tarot card reading online platform with implementation of NLP sentiment analysis(OpenAI API) and Bayesian probability model to provide personalized and context based interpretation based on user queue.
- Implemented real time interactive system reporting through Socket.IO and React.js which allows multiplayer Tarot reading and card interpretation, resulting 50% user satisfaction and 25% reduction in the session dropout rate.
- Deployed full stack architecture online using Node.js on AWS, employed Elastic Load Balancing (ELB) to distribute user traffic across multiple instances of the application and reducing downtime, improving API response time by 20%.

## SKILLS & INTERESTS

**Frontend:** JavaScript, TypeScript, HTML, CSS, Angular, React Native, Webpack, Bootstrap, jQuery

**Backend:** Python (NumPy, Pandas, SciPy), R (Bioconductor, ggplot2, tidyverse), NodeJS, Java, JUnit, REST

**Database:** MySQL, MongoDB, SQLAlchemy, Power BI, Redis, Flyway, Data Modeling, SQLite

**Machine Learning:** OpenAI API, PyMC3, NLP, Bayesian Inference, TensorFlow, PyTorch, Scikit-learn, Deep Learning