Hao-Li Huang

hao-li.huang@yale.edu • GitHub & LinkedIn: haolihuang

Education

Yale University — New Haven, CT
Ph.D. Physical Chemistry (GPA: 3.96/4.00) — *August 2021*

- Taiwan Government Scholarship, one of 125 recipients nationwide

National Taiwan University (Ranks 1st in Taiwan) — Taipei, Taiwan
B.S. Chemistry (GPA: 3.94/4.00) — *June 2013*

- Top 5% of the 70-student class: Presidential Award (twice) & Phi Tau Phi Scholastic Honor Society
- Outstanding Poster Award for undergraduate research

Project management and communication experience

Yale University — New Haven, CT
Graduate Researcher — *June 2016 - Present*

- Identified weak signals from noisy datasets using statistical testing implemented by Python (NumPy, Pandas, SciPy, matplotlib) to differentiate similar chlorophyll molecules, leading to a co-first-author publication
- Modeled Markov process time series to construct complex enzyme kinetic models, and implemented the model using MATLAB to analyze experimental data, leading to new mechanistic insights and a first-author publication
- Mentored an undergraduate student on senior research project through weekly meetings on experimentation, programming, and writing, resulting in a publication
- Collaborated with cross-functional teams, from chemists and biologists to glassblowers and mechanists on 5+ projects; regularly presented summary of data analysis verbally or in writing to stakeholders and interdisciplinary audiences
- Authored 6+ publications; presented posters in regional and international conferences

Yale University — New Haven, CT
Teaching Fellow, 4 Semesters — *September 2015 – May 2017*

- Led weekly discussion sections for 10-16 students; prepared course materials to clarify complex concepts; catalyzed productive group discussions; advised students individually; increased engagement by 20%

Academia Sinica (National Academy of Taiwan), Institute of Atomic and Molecular Sciences — Taipei, Taiwan
Research Assistant — *July 2014 - June 2015*

- Designed innovative apparatuses to enable measurements on short-lived molecules; wrote MATLAB script to efficiently process spectral data; published a first-author paper cited 120+ times

Data science projects (More on https://github.com/haolihuang)

Predictive Modeling & Recommendation System

- Trained Random Forest and XGBoost Classifiers to predict if property maintenance fines in Detroit would be paid on time; built pipeline and customized transformation class to validate models without data leakage (scikit-learn, xgboost)
- Built Support Vector Machine and Logistic Regression models to predict if text message is spam (scikit-learn)
- Built recommendation system on Apache Spark (PySpark, MLlib)

Data Analysis

- Built linear regression model (statsmodels) to analyze multi-dimensional Covid-19 dataset and translated insights into recommendations; cleaned and combined data from multiple sources (NumPy, pandas, RegEx)

Visualization

- Created Tableau dashboard to interactively visualize factors correlating to on-time property maintenance fine payments
- Visualized median housing sale prices in New York City on interactive map (plotly)

Technical skills

- **Programming Languages:** Python (NumPy, pandas, scikit-learn, RegEx, SciPy, statsmodels, NLTK, NetworkX), R (rjags), MATLAB, Relational Database (SQL), Spark (PySpark, MLlib), TensorFlow
- **Visualization Tools:** Tableau, Python (matplotlib, seaborn, plotly)
- **Statistical Tools:** Bayesian statistics, Markov chain Monte Carlo, Numerical simulation, Bootstrapping, A/B testing
- **Coursework:** Probability and Statistics, Linear Algebra, Machine Learning, Text Mining, Social Network Analysis

Activities and interests

- Board games (collection of 60+), Singing (8+ years choral experience), Travel (10+ countries visited)