Shiyu Ma

424-535-8421 | shiyu.ma2023@gmail.com | Los Angeles, CA
github.com/Sylviama1026 | linkedin.com/in/sylvia-ma-2023/

EDUCATION

UNIVERSITY OF CALIFORNIA, LOS ANGELES (UCLA)   Los Angeles, CA
B.S. in Mathematics / Economics and Statistics, Specialization in Computing   Expected June 2023
Cumulative GPA: 3.97/4.0, Deans Honors List(All Quarters)
Relevant Coursework: Experimental Design, Linear Algebra, Probability, Statistical Models and Data Mining, Linear Models, Data Analysis and Regression, Stochastic Processes, Optimization, Statistical Programming with R

SKILLS

SQL, Python(Pandas, NumPy, Matplotlib, Scikit Learn, Plotly, TensorFlow), R, Excel(vlookup), C++, Tableau, Linux
AB Testing, ANOVA, Logistic Regression, KNN, LASSO, Random Forest, Gradient Boosting, Text Mining, SHAP

WORK EXPERIENCE

Thumbtack   San Francisco, CA
Data Analyst Intern   Jun. 2022 - Present
- Built **Random Forest** and **Logistic Model** end-to-end to predict sales conversion probability of prospective users
- Led Feature Selection and Engineering to cut variables used in the model from 52 to 5 with 93% in recall
- Drove model-based strategy cross-functionally with the sales team to prioritize users who are most likely to join
- Designed metrics, **Tableau** dashboard, and power analysis to verify 8% revenue increases in marketing **AB testing**

TAL Education Group   Beijing, China
Data Analyst Intern   Jun. 2021 - Sep. 2021
- Conducted Click Sequences **Clustering** to identify new users' needs by **Python**, raised 5% conversion rate
- Designed revenue, cost, and satisfaction metrics using **SQL** and **Excel** to optimize product lines, cut costs by 8%
- Communicated with engineers and operation teams to refine metrics dimensions and generate weekly reports

Onomy   Los Angeles, CA
Data Analyst Intern   Mar. 2021 - Jun. 2021
- Scraped 47,812 Reddit posts for **text mining** to understand adulting concerns and related business chances
- Inspected **emotion composition** of 8 keywords by **Python** with radar plot, specified marketing tone and strategies
- Performed **topic modeling** for 8 keywords to target specific needs of user subgroups, reflected in a 20-pages report

RESEARCH & PROJECTS

Junction of Statistics and Biology Lab   Los Angeles, CA
Research Assistant   Mar. 2022 - Present
- Built an **R package** to optimize parameters in the Generalized Additive Model that regress on gene expression data
- Constructed **hierarchy trees** for cell types using PCA and various classifiers to refine clustering result resolutions

PARIS Lab   Los Angeles, CA
Research Assistant   Feb. 2022 - Present
- Performed interpretable Symbolic Regression to predict materials' fracture energy, replaced the neutral network
- Conducted SHAP analysis to identify error sources while refining feature designs, increased accuracy by 20%

The Impact of COVID-19 on Global Education   Los Angeles, CA
Team member   Apr. 2021 - Jun. 2021
- Drew **time-series** plot and interactive maps as dashboard in Python to show public awareness changes in COVID
- Created interactive bubble plot in **Plotly** to explore the trade-off between GDP, School Closure Length, and Cases

LEADERSHIP EXPERIENCE & REWARDS

Chinese Students and Scholars Association, *VP of Career Development, 2019-2022*

Queen's Road Foundation Undergraduate Research Fellowship, *1 Recipient per Department, 2022*

ASA DataFest at UCLA, *Finalist, 2022*

AT&T Marketing Strategy Case Competition, *The 3rd Place Finalist, 2020*