Siming, Su (Martin)

Cell: 805-280-9899 **Email**: siming@ucsb.edu

Education

University of California, Santa Barbara September 2018 – September 2022

- Double Major: B.S. Mathematics, B.S. Data Science
- GPA：3.90
- Related Courses: Bayesian Data Analysis, Stochastic Analysis, SAS Programing, Data Science, Regression Analysis, Design of Experiment, Time Series, Machine Learning, Decision Theory.

Internship and related work experience

MINISO (Data Science Intern) | *Python, Excel, MySQL, Tableau* August 2020 – October 2020

- Co-developed a recommendation system using Python for specific segments of a customer bases which boosted 15% of sales the following month.
- Cleaned, prepped the datasets using SQL, Python, and Excel for data scientist leaders to build models that resulted in an increase of the order output by 12%.
- Trained a time series model to predict the accumulated number of VIP members and product sales for the whole company within 95% prediction interval.
- Used Tableau to create weekly data visualization report that increased efficiency for other departments.

Undergraduate Learning Assistant (Bayesian Inference) | *R, Stan* September 2021 – Present

- Coached over 20 students each week in person to achieve success in Bayesian Inference course.
- Collaborated with both teaching staff 2-3 times a week to discuss students' feedback and improvement.
- Created homework, midterm, and practice problems for students to practice the teaching materials.

Research Experience and Related Projects

Guided Research in Deep Learning | *Python, TensorFlow, NumPy* July 2021 – September 2021

- Applied CycleGAN algorithm to generate the pictures of cities filled with garbage using unpaired datasets.
- Collaborated with Professor Tomoyuki Ichiba for the optimization and training, which results in converging discriminator loss of 0.1 and generator loss of 4.69.
- Corresponded with Chinese government officials distributed pictures for propaganda to improve environmental awareness.
- Presented comprehensive research report to Professor Tomoyuki and Chinese local government officials.

Dating App Analysis (link) | *R markdown, Keras, ggplot* January 2022 – March 2022

- Collaborated in a team of 2 to clean the dataset with 10,000 rows from a dating app called Lovoo.
- Explored the data by performing data visualization in various forms including heatmap, bar charts, histograms.
- Predicted the count of likes of users using machine learning algorithms including Linear Regression, XGboost, and Neural Network, achieved over 80% R squared value by choosing the optimal model.
- Compared machine learning algorithms' performances and generated the user report as a web.

Seattle Public Library Data Visualization (link) | *Java, SQL, HTML* January 2022 – March 2022

- Performed Data Mining and used SQL to retrieve and extract data from Seattle Public Library database.
- Cleaned and sorted the data to extract the information of the most popular books in the library.
- Created 3D and 4D data visualizations on the target dataset by using Processing (Java-based).
- Presented comprehensive data visualization report to the visual lab officials.

Skills

- **Programing:** Python (NumPy, Pandas, Scikit-learn, TensorFlow, PyTorch), R (RStudio), C++, Java, SQL, HTML
- **Professional Skills:** Data Visualization, Data Analysis, Regressions, Machine Learning, Deep Learning