Chi (Yuki) Yu

22 Statuary • Irvine, CA 92620 • yuki620@g.ucla.edu • +1 (949) 299-7267

Education

University of California, Los Angeles — **Los Angeles, CA**
B.S., Double Major in Mathematics/Applied and Statistics & Data Science (GPA: 3.88) — **Oct 2020 — Jun 2024**
Honors/Awards: The Big Bang Theory Scholarship
Relevant Coursework: Data Analysis & Regression, Design & Analysis of Experiment, Probability, Mathematical Statistics, Python, C++ Optimization, Monte Carlo Methods, Statistical Consulting, Statistical Models & Data Mining, Stochastic Processes, Game Theory, Accounting

Skills

Technical: Python, R, SQL, SAS, C++, MATLAB, Java, Tableau, Microsoft Suite (Word, Excel, PowerPoint)
Language: English (native), Mandarin Chinese (native), Japanese (Advanced)

Experience

Academia Sinica — **Taipei, Taiwan**
Research Intern — **Jun 2023 — Aug 2023**

- Checked the validity of a mathematical model of biofilm oscillations presented in a research paper and modified it
- Wrote new ODE models and solved them in Python using RK4 and ran simulations in NetLogo to see how the models would behave as individual-based models

Mechanics Bank Auto Finance — **Irvine, CA**
Auto Risk Analyst — **Sep 2022 — Mar 2023**

- Worked in the originations scorecard development team to build a higher accuracy scorecard to minimize the overallocation of resources for risk. This would have freed up resources that could be used towards the current expansion of the bank.
- Pulled, cleaned, and prepared the data to build the development database for a new custom originations scorecard using SAS, SQL, and Excel
- Built a preliminary logistic inference model for the scorecard with SQL and SAS
- Built a finalized version of an inference model that will generate loss predictions with more accuracy Mechanics Bank Auto Finance

Mechanics Bank Auto Finance — **Irvine, CA**
Auto Risk Analyst Intern — **Jun 2022 — Sep 2022**

- Processed data and attributes to prepare the development database for a new custom originations scorecard.
- Analyzed loan default data to determine the optimal variables for the scorecard using SAS, SQL, and Excel.
- Conducted stability testing for the new version of PointPredictive fraud scores by coding macros in SAS to calculate PSI and creating new distributions with real data with SQL.
- Set up and conducted a new quarterly monitoring process for the loss forecast model, including backtesting, stability testing, and stress testing using SAS, SQL, and Excel.
- Created custom economic scenarios for stress testing by stressing HPI, PTI, and UsedCarIndex.

Honors & Awards

ASA DataFest at UCLA 2023 – *Judge's Choice Winner* — **Apr 2023**

- 48-hour data analytics hackathon with almost undergraduates and 80 teams. This year's data was on ABA's Pro Bono Service which provides free legal counseling. Our team won the Judge's Choice Award for humanizing the data
- Our team focused on the responsiveness of this service and its client retention rate with added insight from sentiment analysis.
- Used R, Python, and Tableau to analyze and visualize the data.

Relevant Projects

IMDB Reviews Sentiment Analysis — **Nov 2023**

- Used Python to do NLP and conduct sentiment analysis on the movie reviews left on the IMDB website
- Examined various machine learning algorithms such as logistic regression, logistic regression with SGD, KNN, Random Forest, and Decision Trees under both BoW and TF-IDF

Amazon Price Tracking Bot — **Mar 2023**

- Used Python to create a bot that checks the price of a product at 8:00 a.m. every Monday, Wednesday, Friday, and Saturday.
- If the price drops below our target price, then the bot sends an email to notify the user of the price drop.
- Used BeautifulSoup for web scraping, the SMTP module to send the emails, and crontab to keep the script active.

Kaggle Project: Predicting Winning Proportions — **Jan 2022 — Mar 2022**

- Used R to create the best model for predicting winning proportions with college basketball data.
- Analyzed the data to determine the significance of the 21 predictors, how well these predictors described the data, and the necessary transformations to create the best possible predictive model.