

Gyujin Seo

910 Weyburn Place, Los Angeles, CA, 90024
seokj98@g.ucla.edu | <http://gyujinseo.github.io>
(310)-869-4903

EDUCATION

June 2023 UNIVERSITY OF CALIFORNIA, LOS ANGELES Los Angeles, CA
(Expected)
Bachelor of Science in Statistics and Data Science(Senior year)

Relevant Coursework: ●Statistical Programming with R, ●Probability, ●Mathematical Statistics, ●Data Analysis and Regression, ●Computational Statistics with R, ●Linear Models, ●Design and Analysis of Experiment, ●Computation and Optimization for Statistics, ●Statistical Models and Data Mining, ●Monte Carlo Methods, ●Practice of Statistical Consulting, ●Pattern Recognition and Machine Learning, ●Statistical Models in Finance

CORE COMPETENCIES

- **Programming:** R, Python
- **Visualization:** Tableau Software
- **Version Control:** Git

Work Experience

July 2022 - **PREDICTING CUSTOMER CHURN**
August 2022

- For the stability of the company's financial problems, create a model to predict what customer has a high risk of leaving the bank's credit card services using R
- Compared model performance of various machine learning algorithms and regression models to find the best model
- Random Forest has an AUC score of 98% that shows the model is learning the data well enough.

PROJECTS

August 2020 **COVID-19 LOCKDOWN ANALYSIS**

- Analyzed how lockdown has affected the Covid-19
- Made visualizations to analyze using Tableau Software
- Oversaw the cleaning, refining, augmenting of the dataset, ensuring standardized, usable, and well-formatted data.
- Lockdown was found to be effective in lowering the number of COVID-19 cases in all three countries, as there was a consistent pattern of decreasing cases during and after the lockdown.

December 2022 **Predicting the Useful Values with Machine Learning Models**

- This project aimed to identify the best machine learning model for predicting whether a review on Yelp is "useful." using R.
- Pre-processed the data by removing irrelevant variables, correcting errors, and converting the text-based reviews into numerical variables using NLP
- Experimented with several ML models, and concluded that the Random Forest Model performed the best based on the ROC curve and AUC.

EXTRA ACTIVITIES

● KSEA(Korean-American Scientists and Engineers Association) : As a member of KSEA, I made significant contributions to the organization's efforts in hosting various social and academic events and fostering active exchange among students majoring in STEM fields.