Gyujin Seo

107, Hwagok-ro 13-gil, Gangseo-gu, Seoul, 07703, Republic of Korea
seokj98@g.ucla.edu| http://gyujinseo.github.io
+82 10-4013-2898

EDUCATION

June 2023 **University of California, los angeles** Los Angeles, CA

Bachelor of Science in Statistics and Data Science

Honors: Dean's Honors List

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
- **Database**: MySQL
- **Visualization**: Tableau Software
- **Version Control**: Git

PROJECTS

July 2020 - 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.

July 2022 - August 2022 **Predicting customer churn**

- 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.

October 2022 - 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.