Richard Lim

201-820-8219 | richardlim7100@gmail.com | Web Portfolio | Waldwick, NJ

Education

Villanova University — May 2024

B.S. in Statistics; Minor in Mathematics / GPA: 3.52

Coursework: Data Science, Algorithms and Data Structures I and II, Applied Statistical Models, Data Base Systems, Linear Algebra, Design of Statistical Experiments, Mathematical Statistics I, Discrete Mathematics, Calculus I, II, and III

Experience

Villanova Baseball — **Villanova, PA**

Data Analyst — February 2022 - Present

- **Improved team record from 3-13-1 to 16-16** by producing in-depth scouting reports with data analyses on opposing teams' pitching and hitting tendencies.
- **Analyzed opponents' data** utilizing R and Synergy Baseball. Communicated findings and analyses with analyst team.
- **Presented data analyses to coaches** using strong written and verbal communication skills. Developed in-game and developmental strategies with coaches, using analyses.

Villanova Recreation — **Villanova, PA**

Fitness Center Receptionist — August 2021 - Present

- **Reduced damages by 95% and accidents by 98%** by providing comprehensive instructions to over 50 daily facility users on the proper and safe operation of equipment.
- Consistently enforced all facility and program policies **resulting in a 100% adherence to policies.**

CoSIBS — **Denver, CO**

Fellow — June 2022 – August 2022

- **Built a random forest model with an out-of-bag error rate of 3.87:** Implemented random forest machine learning to a large dataset and identified predictors of obesity in American youth.
- **Developed presentation given to non-technical audience** by using data wrangling and data visualization techniques. Produced models of geospatial distribution, decision trees, and partial dependence plots.
- **Finished top 5 in a collaborative 6-hour long hackathon**. Implemented machine learning models, including random forest and linear regression to predict the duration of intubation. Presented analysis to panel of judges.

Projects

Predicting Survival of Passengers on the Titanic | Python / Kaggle

Performed exploratory data analysis, feature engineering, scaling, and created custom transformers and pipeline on Titanic data set to create model that predicts passenger survivability. Trained a Random Forest model and used hyperparameter tuning to find best possible model. The model was given a score of 83%.

Breast Cancer Classification using Neural Networks | Python / Kaggle

Conducted data loading, preprocessing, and feature standardization on the Breast Cancer Wisconsin dataset. Constructed a neural network model architecture using Keras Sequential API, optimizing it with Adam optimizer and sparse categorical cross-entropy loss. Trained the model with training data and evaluated its accuracy on test data, achieving an accuracy of 96%

Activities

Sports Analytics Club, Villanova University — August 2021 - Present

Club Sailing, TIDE Representative, Villanova University — August 2021 - Present

- **Increased participation of minority groups by 500%** by leading team in discourses on diversity, equity, and inclusion and how to present themselves in a predominantly Caucasian sport.

Villanova Leadership Program, Villanova University — January 2022 - April 2022

Skills

Programming and Database: Python (Matplotlib, Seaborn, Sklearn), R (dplyr, tidyr, ggplot2, Shiny), SQL, Java, HTML

Analysis Techniques: Neural Networks, Logistic Regressions, Decision Trees, Random Forests, Hypothesis Testing

Tools: Tableau, Kaggle, R Markdown, Jupyter, PyCharm, RStudio, VSCode, Excel (Power Query, V-Look Up, Pivot Tables)

Languages: Fluent in English and Korean, proficient in Spanish